

STIC-Biotech/ChemLib

106866

From: Li, Ruixiang
Sent: Monday, October 27, 2003 5:12 PM
To: STIC-Biotech/ChemLib
Subject: Sequence search of Application NO:10/070,241

Please do a standard search on

- (i) SEQ ID NO: 2 against commercial nucleic acid databases;
- (ii) SEQ ID NO: 1 against commercial amino acid databases.

Thank you very much!

Ruixiang Li
GAU 1646
CM1 10E18
Mail Box 10D19
306-0282

10/28/03
10/28/03 4:00 PM
STIC-Biotech/ChemLib

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 10/28/03
Date Completed: 10/28/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH: /
NA Sequences: _____
AA Sequences: / _____
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: 01/02
Sequence Sys.: _____
WWW/Internet: _____
Other (specify): _____

```
RESULT 6
US-10-272-983-37
; Sequence 37, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huang T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/272,983
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 37
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-272-983-37

Query Match 53.6%; Score 1294.4; DB 12; Length 1296;
Best Local Similarity 99.9%; Pred. No. 2.9e-285;
Matches 1295; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 354 ATGACGGCGCTTAACATTACCCCGGAGCAGTCTCTCGGCTGCTGGGACCAACCTG 413
DB 1 ATGACGGCGCTTAACATTACCCCGGAGCAGTCTCTCGGCTGCTGGGACCAACCTG 60

QY 414 ACGCGGAGCAGTTCATCGCTCTGTACCGGCTGCGACCGCTCTGTACACCCGAGAGCTG 473
DB 61 ACGCGGAGCAGTTCATCGCTCTGTACCGGCTGCGACCGCTCTGTACACCCGAGAGCTG 120

QY 474 CCGGACGGCCCAAGCTGCTGTCTACCGGCTGCTCATCTTCCCTCGGCTC 533
DB 121 CCGGACGGCCCAAGCTGCTGTCTACCGGCTGCTCATCTTCCCTCGGCTC 180

QY 534 TTGGCAATGCTCTGGTGTCTACGTGGTGACCCGACGAGCCATGCGACCCCTCAC 593
DB 131 TTGGCAATGCTCTGGTGTCTACGTGGTGACCCGACGAGCCATGCGACCCCTCAC 240

QY 594 AACATCTTTATCTGCTCTGCGGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATTC 653
DB 241 AACATCTTTATCTGCTCTGCGGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATTC 300

QY 654 GTCACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGGTCTTTTCATTTGCAAGATG 713
DB 301 GTCACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGGTCTTTTCATTTGCAAGATG 360

QY 714 GTGCCATTGTCCAGTCTACCGCTGTGTGACAGAACTCTCACTATGACCTGCAATGCT 773
DB 1 GTGCCATTGTCCAGTCTACCGCTGTGTGACAGAACTCTCACTATGACCTGCAATGCT 420

361 GTGCCATTGTCCAGTCTACCGCTGTGTGACAGAACTCTCACTATGACCTGCAATGCT 420
QY 774 GTGGAAGGCACACGAGGACTTGTGCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 833
DB 421 GTGGAAGGCACACGAGGACTTGTGCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 480
QY 834 AGGGCTTTTCACAATGCTAGGTGTGTCTGGCTGGTGGGAGTCACTGTAGATCAACCATG 893
DB 481 AGGGCTTTTCACAATGCTAGGTGTGTCTGGCTGGTGGGAGTCACTGTAGATCAACCATG 540
QY 894 TGGCACGTGCAACAACTTGGAGATCAAAATGATGCTTCTTATATGAAAGGAAACACATCTGC 953
DB 541 TGGCACGTGCAACAACTTGGAGATCAAAATGATGCTTCTTATATGAAAGGAAACACATCTGC 600
QY 954 TGCCTTAGAAGAGTGGACGAGCCCTGTGCACCAAGAGATCAACCACTTCACTCTGTTC 1013
DB 601 TGCCTTAGAAGAGTGGACGAGCCCTGTGCACCAAGAGATCAACCACTTCACTCTGTTC 660
QY 1014 ATCCCTCTTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 1073
DB 661 ATCCCTCTTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 720
QY 1074 CTTTGGATAAAGAAAGAGTTGGGGATGCTTCAGTGTCTTGGAACTATTTCATGAAAGAA 1133
DB 721 CTTTGGATAAAGAAAGAGTTGGGGATGCTTCAGTGTCTTGGAACTATTTCATGAAAGAA 780
QY 1134 ATGTCACAAATAGCCAGGAAGAAAGAGCTGTCTATGATGCTGACATGCTGGTGGCT 1193
DB 781 ATGTCACAAATAGCCAGGAAGAAAGAGCTGTCTATGATGCTGACATGCTGGTGGCT 840
QY 1194 CTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1253
DB 841 CTCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
QY 1254 TTTGAAAAGGAATATGATGATGTCACAATCAAGATGATTTTGTGTGTGTGTGTGTGTGT 1313
DB 901 TTTGAAAAGGAATATGATGATGTCACAATCAAGATGATTTTGTGTGTGTGTGTGTGTGT 960
QY 1314 GGATTTTCCAACTCCATCTGTAAATCCCAATGCTATGATGATGATGATGATGATGATGATGAT 1373
DB 961 GGATTTTCCAACTCCATCTGTAAATCCCAATGCTATGATGATGATGATGATGATGATGATGAT 1020
QY 1374 AAAATGTTTTGTCTGCAAGTTTGTATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1433
DB 1021 AAAATGTTTTGTCTGCAAGTTTGTATGCAATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1434 AGGCATGGAATTTACGGAATTTACATGATGCGGAAGAAAGCAAAAGTTTTCCTCAGAGAG 1493
DB 1081 AGGCATGGAATTTACGGAATTTACATGATGCGGAAGAAAGCAAAAGTTTTCCTCAGAGAG 1140
QY 1494 AATCCAGTGGAGGAAACCAAGAGGAGAACATTCAGTGATGCGCAACATTCAGTGAATG 1553
DB 1141 AATCCAGTGGAGGAAACCAAGAGGAGAACATTCAGTGATGCGCAACATTCAGTGAATG 1200
QY 1554 TGTGAACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1613
DB 1201 TGTGAACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1614 CTGGCTGAGAAATTCCTTTTACAGAGTGGGGATTA 1649
DB 1261 CTGGCTGAGAAATTCCTTTTACAGAGTGGGGATTA 1296

RESULT 7
US-10-393-807-37
; Sequence 37, Application US/10393807
; Publication No. US20030175891A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huang T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
```

```

/ FILE REFERENCE: AREN00050
/ CURRENT APPLICATION NUMBER: US/10/393,807
/ CURRENT FILING DATE: 2003-03-21
/ PRIOR APPLICATION NUMBER: US/09/417,044
/ PRIOR FILING DATE: 1999-10-12
/ PRIOR APPLICATION NUMBER: 60/109,213
/ PRIOR FILING DATE: 1998-11-20
/ PRIOR APPLICATION NUMBER: 60/120,416
/ PRIOR FILING DATE: 1999-02-16
/ PRIOR APPLICATION NUMBER: 60/121,851
/ PRIOR FILING DATE: 1999-02-26
/ PRIOR APPLICATION NUMBER: 60/123,946
/ PRIOR FILING DATE: 1999-03-12
/ PRIOR APPLICATION NUMBER: 60/123,949
/ PRIOR FILING DATE: 1999-03-12
/ PRIOR APPLICATION NUMBER: 60/136,436
/ PRIOR FILING DATE: 1999-05-28
/ PRIOR APPLICATION NUMBER: 60/136,437
/ PRIOR FILING DATE: 1999-05-28
/ PRIOR APPLICATION NUMBER: 60/136,439
/ PRIOR FILING DATE: 1999-05-28
/ PRIOR APPLICATION NUMBER: 60/136,567
/ PRIOR FILING DATE: 1999-05-28
/ Remaining Prior Application data removed
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 37
/ LENGTH: 1296
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-393-807-37

```

Query Match	53.6%	Score 1294.4	DB 12	Length 1296
Best Local Similarity	99.9%	Pred. No. 2.9e-285		
Matches 1295; Conservative	0	Mismatches 1	Indels 0	Gaps 0
QY	354	ATCGAGCGCTTAA	CATTACCCCGAGCAGTTCTCTCGGCTGCTCGGAGCACAACTGG	413
DB	1	ATCGAGCGCTTAA	CATTACCCCGAGCAGTTCTCTCGGCTGCTCGGAGCACAACTGG	60
QY	414	ACGCGGAGCAGTT	CATCGCTCTGTACCGGTGCGACCGCTCGTACACCCCGACAGCTG	473
DB	61	ACGCGGAGCAGTT	CATCGCTCTGTACCGGTGCGACCGCTCGTACACCCCGACAGCTG	120
QY	474	CCGGGACGCGC	CAAGCTGGCCCTCGTGTCTACCGGCGTGTCTATCTTGGCGCTGGCGCTC	533
DB	121	CCGGGACGCGC	CAAGCTGGCCCTCGTGTCTACCGGCGTGTCTATCTTGGCGCTGGCGCTC	180
QY	534	TTTGGCAATGCT	CTGTGGTGTCTACGTGTGTACCCGCGACAGGCCATCGCACCGTCACC	593
DB	181	TTTGGCAATGCT	CTGTGGTGTCTACGTGTGTACCCGCGACAGGCCATCGCACCGTCACC	240
QY	594	AACATCTTTAT	CTGCTCTTGGCGCTCAGTGACCTGTCTATCACTTCTTGTGCATTTCC	653
DB	241	AACATCTTTAT	CTGCTCTTGGCGCTCAGTGACCTGTCTATCACTTCTTGTGCATTTCC	300
QY	654	GTCAACATGCT	CCAGAACATTTCCGACAACTTCCGGCGTGGGGGGTCTTTCATTATGCAAGATG	713
DB	301	GTCAACATGCT	CCAGAACATTTCCGACAACTTCCGGCGTGGGGGGTCTTTCATTATGCAAGATG	360
QY	714	GTGCCATTTGT	CCAGTCTACCGCTGTTGTGACAGAAATCCCTCACTATGACCTGCATTTGCT	773
DB	361	GTGCCATTTGT	CCAGTCTACCGCTGTTGTGACAGAAATGCTCACTATGACCTGCATTTGCT	420
QY	774	GTGGAAGGCAC	AGGACATTGTGCATCCTTTTAAATGAAATGAGTGGCAATATACCAACCGA	833
DB	421	GTGGAAGGCAC	AGGACATTGTGCATCCTTTTAAATGAAATGAGTGGCAATATACCAACCGA	480
QY	834	AGGCTTTTACA	TGCTAGGTGTGCTGTGGCTGTGGCGAGTCATCGTAGGATCACCCATG	893
DB	481	AGGCTTTTACA	TGCTAGGTGTGCTGTGGCTGTGGCGAGTCATCGTAGGATCACCCATG	540
QY	894	TGGACGTGCA	CAACTTGGAGTCAAAATATGACTTCTTATATGAAAGGAAACACATCTGC	953

```

RESULT 8
US 10-2225-567A-509
    : Sequence 509, Application US/10225567A
    : Publication No. US20030113798A1
    : GENERAL INFORMATION:
    : APPLICANT: LifeSpan Biosciences
    : APPLICANT: Brown, Joseph P.
    : APPLICANT: Burner, Glenn C.
    : APPLICANT: Roush, Christine L.
    : TITLE OF INVENTION: ANTIGENIC PEPTIDE
    : FILE REFERENCE: 1920-4-4
    : CURRENT APPLICATION NUMBER: US/10/2225
    : CURRENT FILING DATE: 2001-12-19
    : PRIOR APPLICATION NUMBER: 60/257,144
    : PRIOR FILING DATE: 2000-12-19
    : NUMBER OF SEQ ID NOS: 2292
    : SOFTWARE: Patent version 3.1
    : SEQ ID NO 509
    : LENGTH: 1368
    : TYPE: DNA
    : ORGANISM: Homo sapiens

```

```
QY 421 LAENSPDSGH 431
DB 421 LAENSPDSGH 431
RESULT 4
US-10-272-983-38
; Sequence 38, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
; APPLICANT: Dang, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/272,983
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1998-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 74
; SEQ ID NO 38
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-272-983-38
Query Match 99.9%; Score 2232; DB 12; Length 431;
Best Local Similarity 99.8%; Pred. No. 3.2e-194;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQALNITPEQSRLLRDHNLTRQFIYLRPLVYTPPELPGRAKLALVLTGVLFALAL 60
DB 1 MQALNITPEQSRLLRDHNLTRQFIYLRPLVYTPPELPGRAKLALVLTGVLFALAL 60
QY 61 FGNALVFVVTTRSKAMRTVTNIFCSLSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
DB 61 FGNALVFVVTTRSKAMRTVTNIFCSLSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
QY 121 VPVQSTAVVTETLTMTCTIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVWLVAVTVGSPM 180
DB 121 VPVQSTAVVTETLTMTCTIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVWLVAVTVGSPM 180
QY 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240
DB 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240
QY 241 LWIKRVDGSLRTIHGKEMSKIARKKRAVIMVTVVAVFVAVCWAPFVHVHMMIEYSN 300
DB 241 LWIKRVDGSLRTIHGKEMSKIARKKRAVIMVTVVAVFVAVCWAPFVHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFNFKNVL SAVCYCVNKTSPAQ 360
DB 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFNFKNVL SAVCYCVNKTSPAQ 360
US-10-393-807-38
; Sequence 38, Application US/10393807
; Publication No. US20030175891A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/393,807
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 74
; SEQ ID NO 38
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-393-807-38
Query Match 99.9%; Score 2232; DB 12; Length 431;
Best Local Similarity 99.8%; Pred. No. 3.2e-194;
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQALNITPEQSRLLRDHNLTRQFIYLRPLVYTPPELPGRAKLALVLTGVLFALAL 60
DB 1 MQALNITPEQSRLLRDHNLTRQFIYLRPLVYTPPELPGRAKLALVLTGVLFALAL 60
QY 61 FGNALVFVVTTRSKAMRTVTNIFCSLSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
DB 61 FGNALVFVVTTRSKAMRTVTNIFCSLSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
QY 121 VPVQSTAVVTETLTMTCTIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVWLVAVTVGSPM 180
DB 121 VPVQSTAVVTETLTMTCTIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVWLVAVTVGSPM 180
QY 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240
DB 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYTTFTILVILFLLPLMVLILYSKIGYE 240
QY 241 LWIKRVDGSLRTIHGKEMSKIARKKRAVIMVTVVAVFVAVCWAPFVHVHMMIEYSN 300
DB 241 LWIKRVDGSLRTIHGKEMSKIARKKRAVIMVTVVAVFVAVCWAPFVHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFNFKNVL SAVCYCVNKTSPAQ 360
DB 301 FEKEYDDVTIKMIFAIVQIIGFSNSICNPVYAFMNFNFKNVL SAVCYCVNKTSPAQ 360
```


QY 241 LWKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 300
 DB 241 LWKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 300
 QY 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 DB 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 QY 361 RHNSGITMMRKAFLSRENPEETKGEAFSDGNIEVKLCEOTEKKKLRHLALFRSE 420
 DB 361 RHNSGITMMRKAFLSRENPEETKGEAFSDGNIEVKLCEOTEKKKLRHLALFRSE 420
 QY 421 LAENSPDLSGH 431
 DB 421 LAENSPDLSGH 431

RESULT 6
 US-10-080-263C-2
 ; Sequence 2, Application US/10080263C
 ; Publication No. US20030143670A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bonini, James A.
 ; APPLICANT: Huang, Ling-Yan
 ; APPLICANT: Wilson, Amy
 ; TITLE OF INVENTION: DNA ENCODING SNORF44 RECEPTOR
 ; FILE REFERENCE: 1795/59370-A/JPW/ADM/ANX
 ; CURRENT APPLICATION NUMBER: US/10/080,263C
 ; CURRENT FILING DATE: 2002-02-20
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 415
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-080-263C-2

Query Match 95.5%; Score 2135; DB 12; Length 415;
 Best Local Similarity 96.3%; Pred. No. 2e-185;
 Matches 415; Conservative 0; Mismatches 0; Indels 16; Gaps 1;
 QY 1 MQALNITPQFSLRLDHNLTREOFIALYRLPLVVTPELPGRKALALVLTGVLI FALAL 60
 DB 1 MQALNITPQFSLRLDHNLTREOFIALYRLPLVVTPELPGRKALALVLTGVLI FALAL 44
 QY 61 FGNALVFYVVTTRSKAMRTVTNIFICSLSLSDLLITFCIPVTMLQNSDNWLGGA FICKM 120
 DB 45 FGNALVFYVVTTRSKAMRTVTNIFICSLSLSDLLITFCIPVTMLQNSDNWLGGA FICKM 104
 QY 121 VPFVOSTAVVTEILMTCTIAVERHOGVLHPFKMKQYTNRRAPTMVGVLVAVI VGS PM 180
 DB 105 VPFVOSTAVVTEILMTCTIAVERHOGVLHPFKMKQYTNRRAPTMVGVLVAVI VGS PM 164
 QY 181 WHVQOLEIKYDFLYEKEHICCLLEWTSPPHQKIYTTFFILVILPLLMWMLILY SKIGYE 240
 DB 165 WHVQOLEIKYDFLYEKEHICCLLEWTSPPHQKIYTTFFILVILPLLMWMLILY SKIGYE 224
 QY 241 LMKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 300
 DB 225 LMKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 284
 QY 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 DB 285 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 344
 QY 361 RHNSGITMMRKAFLSRENPEETKGEAFSDGNIEVKLCEOTEKKKLRHLALFRSE 420
 DB 345 RHNSGITMMRKAFLSRENPEETKGEAFSDGNIEVKLCEOTEKKKLRHLALFRSE 404
 QY 421 LAENSPDLSGH 431
 DB 405 LAENSPDLSGH 415

RESULT 7
 US-09-990-940-16
 ; Sequence 16, Application US/09990940
 ; Publication No. US20030027252A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Tian, Hui
 ; APPLICANT: Zhao, Jiagang
 ; APPLICANT: Chen, Jin-Long
 ; APPLICANT: Cutler, Gene
 ; APPLICANT: An, Songzhu
 ; APPLICANT: Dai, Kang
 ; APPLICANT: Gupte, Jamila S.
 ; APPLICANT: Tulazik Inc.
 ; TITLE OF INVENTION: NO. US20030027252A1e1 Receptors
 ; FILE REFERENCE: 018781-007410US
 ; CURRENT APPLICATION NUMBER: US/09/990,940
 ; CURRENT FILING DATE: 2001-11-21
 ; PRIOR APPLICATION NUMBER: US 60/252,841
 ; PRIOR FILING DATE: 2000-11-22
 ; PRIOR APPLICATION NUMBER: US 60/257,636
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: US 60/261,377
 ; PRIOR FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: US 60/279,554
 ; PRIOR FILING DATE: 2001-03-28
 ; PRIOR APPLICATION NUMBER: US 60/280,696
 ; PRIOR FILING DATE: 2001-03-29
 ; NUMBER OF SEQ ID NOS: 54
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 16
 ; LENGTH: 433
 ; TYPE: PRT
 ; ORGANISM: Mus sp.
 ; FEATURE:
 ; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346a
 US-09-990-940-16

Query Match 85.6%; Score 1914; DB 11; Length 433;
 Best Local Similarity 83.5%; Pred. No. 0.4e-165;
 Matches 360; Conservative 27; Mismatches 44; Indels 0; Gaps 0;
 QY 1 MQALNITPQFSLRLDHNLTREOFIALYRLPLVVTPELPGRKALALVLTGVLI FALAL 60
 DB 1 MQALNITPQFSLRLDHNLTREOFIALYRLPLVVTPELPGRKALALVLTGVLI FALAL 60
 QY 61 FGNALVFYVVTTRSKAMRTVTNIFICSLSLSDLLITFCIPVTMLQNSDNWLGGA FICKM 120
 DB 61 FGNALVFYVVTTRSKAMRTVTNIFICSLSLSDLLITFCIPVTMLQNSDNWLGGA FICKM 120
 QY 121 VPFVOSTAVVTEILMTCTIAVERHOGVLHPFKMKQYTNRRAPTMVGVLVAVI VGS PM 180
 DB 121 VPFVOSTAVVTEILMTCTIAVERHOGVLHPFKMKQYTNRRAPTMVGVLVAVI VGS PM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLLEWTSPPHQKIYTTFFILVILPLLMWMLILY SKIGYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICCLLEWTSPPHQKIYTTFFILVILPLLMWMLILY SKIGYE 240
 QY 241 LMKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 300
 DB 241 LMKRRVGGSVLRTIHGKMSKIARKKRAVIMMVTVALFAVCWAPFVHMMIEYSN 300
 QY 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 DB 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 QY 361 RHNSGITMMRKAFLSRENPEETKGEAFSDGNIEVKLCEOTEKKKLRHLALFRSE 420
 DB 361 KPGNSGIMMOKRAKLSRQRPVABEAKGDLFSDANDVVKLCEQPKQKQLKLAFFSSE 420
 QY 421 LAENSPDLSGH 431
 DB 421 LAENSPDLSGH 431

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 10:08:35 ; Search time 86 seconds
(without alignments)
795.479 Million cell updates/sec

Title: US-10-070-241B-1

Perfect score: 2235

Sequence: 1 MQALNTPQFRLRDHNL.....RHLAIFRSLAENSLDSGH 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

```

1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	100.0	431	21	AA094993
2	2235	100.0	431	22	AA064289
3	2235	100.0	431	22	AA074773
4	2235	100.0	431	22	AA048963
5	2235	100.0	431	23	AA097735
6	2235	100.0	431	23	AA097735
7	2235	100.0	431	24	AB073503
8	2232	99.9	431	21	AA071309
9	2232	99.9	431	21	AA028443

10	2226	99.6	431	21	AA028553	Human G protein co
11	1914	85.6	433	23	AAU97740	Mouse TGR346a poly
12	1767	79.1	416	23	AAU97741	Mouse TGR346b poly
13	1288	56.7	245	24	ABP82011	Human G protein-co
14	508	22.7	417	21	AA079377	Rat neuropeptide F
15	502	22.5	420	20	AAW81358	Human 7-transmembr
16	502	22.5	420	21	AA099930	HLWAR77 polypeptid
17	502	22.5	420	21	AA079375	Human neuropeptide
18	502	22.5	420	22	AAU00233	Orphan receptor HG
19	502	22.5	420	23	AAU79013	Human G protein-co
20	502	22.5	522	22	ABE11723	Human GPCR homolog
21	501.5	22.4	420	21	AAU07426	Amino acid sequenc
22	501.5	22.4	522	22	AAU00234	Orphan receptor HG
23	501.5	22.4	522	24	ABP82012	Human neuropeptide
24	501.5	22.4	522	24	ABG73522	Human G-protein co
25	496	22.2	408	21	AA078882	Human NPY-Y7 recep
26	493.5	22.1	375	24	ABG73520	D. rerio G-protein
27	479	21.4	381	16	AA078272	Rat hippocampal ne
28	475.5	21.3	405	21	AA076883	Mouse NPY-Y7 recep
29	474	21.2	381	16	AA078273	Rat hippocampal ne
30	469.5	21.0	381	21	AA058888	Human Y2 receptor
31	469.5	21.0	381	22	AAE07921	Human neuropeptide
32	469.5	21.0	381	22	AAE07957	Human neuropeptide
33	469.5	21.0	381	22	AAE06689	Human neuropeptide
34	469.5	21.0	381	22	AAE05109	Human neuropeptide
35	469.5	21.0	381	22	AAE05120	Human neuropeptide
36	469.5	21.0	381	23	ABE04500	Human hippocampus
37	469.5	21.0	381	24	ABP81858	Human neuropeptide
38	469.5	21.0	381	24	ABG73521	Human NY2R protein
39	468.5	21.0	381	19	AAW41710	Rhesus monkey neur
40	464.5	20.8	381	16	AA078271	Human hippocampal
41	459.5	20.6	460	22	AA061970	Rat HCRTR2 polypep
42	459.5	20.6	460	24	ABG73515	Rat OX2R protein S
43	457	20.4	444	22	AAE04416	Amino acid sequenc
44	455	20.4	444	22	AAE08007	Human hypocretin r
45	455	20.4	444	22	AAE061969	Human HCRTR2 polyp

ALIGNMENTS

```

RESULT 1
AA094993
ID AA094993 standard; Protein; 431 AA.
XX
AC AA094993;
XX
XX 19-JUN-2000 (first entry)
XX
DE Human secreted protein vc38_1, SEQ ID NO:26.
DE
XX
XX Human; secreted protein; cancer; tumour; cardiovascular disorder;
XX blood disorder; haemophilia; autoimmune disease; diabetes; inflammation;
XX infection; fungal; bacterial; viral; HIV; allergy; arthritis;
XX neurodegenerative disease; asthma; contraceptive.
XX Homo sapiens.
XX
XX WO200011015-A1.
XX
XX 02-MAR-2000.
XX
XX 24-AUG-1999; 99WO-US19351.
XX
XX 24-AUG-1998; 98US-0097638.
XX 24-AUG-1998; 98US-0097659.
XX 09-SEP-1998; 98US-0099618.
XX 28-SEP-1998; 98US-0102092.
XX 25-NOV-1998; 98US-0109978.
XX 23-DEC-1998; 98US-0113645.
XX 23-DEC-1998; 98US-0113646.
XX 23-AUG-1999; 99US-0379246.
XX

```

PA (ALPH-) ALPHAGENE INC.
 XX Valenzuela D, Yuan O, Hoffman H, Hall J, Rapiejko P;
 PI WPI; 2000-224657/19.
 DR
 XX New secreted or transmembrane proteins and polynucleotides encoding
 PT them, useful for treating neurodegenerative disorders, autoimmune
 PT diseases and cancer -
 XX
 XX Claim 35; Page 284-285; 357pp; English.
 PS
 XX The invention relates to 40 human secreted proteins (AA94981-Y95020),
 CC and cDNA sequences encoding them (AA23423-A23462). The secreted
 CC proteins of the invention include those that are thought to be only
 CC partially secreted, i.e., transmembrane proteins. The proteins of the
 CC invention may exhibit one or more activities selected from the following:
 CC cytokine activity; cell proliferation; differentiation; immune
 CC modulation; haematopoiesis regulation; tissue growth activity;
 CC activin/inhibin activity; chemotactic/chemokinetic activity; haemostatic
 CC and thrombolytic activity; anti-inflammatory activity; and tumour
 CC inhibition activity. The proteins may be administered to patients as
 CC vaccines, and the nucleotides may be used as part of a gene therapy
 CC regime. Diseases or conditions that may be treated using the proteins or
 CC nucleotides of the invention include autoimmune diseases; genetic
 CC disorders; haemophilia; cardiovascular diseases; cancer; bacterial,
 CC fungal and viral infections, especially HIV; multiple sclerosis;
 CC rheumatoid arthritis; pulmonary inflammation; Guillain-Barre syndrome;
 CC insulin dependent diabetes mellitus; and allergic reactions such as
 CC asthma and anaemia. They may also be used for treating wounds, burns,
 CC ulcers, osteoporosis, osteoarthritis, periodontal diseases, Alzheimer's
 CC disease, Parkinson's disease, Huntington's disease and amyotrophic
 CC lateral sclerosis (ALS). Proteins with activin/inhibin activity may
 CC additionally be useful as contraceptives. Nucleic acid sequences of the
 CC invention may be used in chromosome mapping, and as a source of
 CC diagnostic primers and probes. The present sequence represents one of the
 CC 40 proteins of the invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2235; DB 21; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQALNITPEQFSLRRDHNLTREQFIALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 DB 1 MQALNITPEQFSLRRDHNLTREQFIALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNLGGAFICKM 120
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNLGGAFICKM 120
 QY 121 VPVQSTAVVTEILTMTCTIAVERHOGVLVHPFKMKQYTNRRAPFTMLGVVWLVAVIGSPM 180
 DB 121 VPVQSTAVVTEILTMTCTIAVERHOGVLVHPFKMKQYTNRRAPFTMLGVVWLVAVIGSPM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLEBWSVPVHOKIYTTFFILVILFLLPLMWMLILYSKIGYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICCLEBWSVPVHOKIYTTFFILVILFLLPLMWMLILYSKIGYE 240
 QY 241 LWTIKRVGDSVLRTHGKMSKIARKKGAIVTMVTVVALFVACWAPVHVHMMIEYSN 300
 DB 241 LWTIKRVGDSVLRTHGKMSKIARKKGAIVTMVTVVALFVACWAPVHVHMMIEYSN 300
 QY 301 FEKEYDDVTIKMFAIVQIIGFNSNICNPVYAFMNFKNVLSAVCYVKNKTFSPAQ 360
 DB 301 FEKEYDDVTIKMFAIVQIIGFNSNICNPVYAFMNFKNVLSAVCYVKNKTFSPAQ 360
 QY 361 RHNGSGITMRRKAKFSLRENPEETKGEAFSDGNIEVKICEOTEKKLKRHLAFRSE 420
 DB 361 RHNGSGITMRRKAKFSLRENPEETKGEAFSDGNIEVKICEOTEKKLKRHLAFRSE 420
 QY 421 LAENSPDLSGH 431

DB 421 LAENSPDLSGH 431
 RESULT 2
 AAG64289
 ID AAG64289 standard; Protein; 431 AA.
 XX
 AC AAG64289;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRV11.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; human; guanosine triphosphate;
 KW G-protein.
 XX
 OS Homo sapiens.
 XX
 FN WO200148189-A1.
 XX
 PD 05-JUL-2001.
 XX
 XX 28-DEC-2000; 2000WO-JP09409.
 XX
 XX 28-DEC-1999; 99JP-0375152.
 PR 31-MAR-2000; 2000JP-0101339.
 PR 23-MAY-2000; 2000JP-0155978.
 XX
 PA (HELI-) HELIX RES INST.
 XX
 PI Matsumoto S, Oda T, Saito Y, Morikawa N, Yoshida K, Suwa M;
 PI Sugiyama T;
 XX
 WPI; 2001-425663/45.
 DR N-PSDB; AAA49502.
 XX
 PT Family of guanosine triphosphate binding protein coupled receptors and
 PT genes encoding them for treatment and prevention of diseases associated
 PT with these receptors -
 XX
 PS Claim 1; Pages 59-62; 137pp; Japanese.
 XX
 CC The present sequence is the protein sequence for a human guanosine
 CC triphosphate (GTP)-binding protein-coupled receptor. The receptor is
 CC useful for the investigation, diagnosis, treatment and prevention of
 CC diseases associated with GTP-binding protein-coupled receptors, including
 CC neurological, circulatory, digestive system, immune system, muscle and
 CC urinary system disorders. GTP-binding proteins are also known as
 CC G-proteins.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2235; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MQALNITPEQFSLRRDHNLTREQFIALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 DB 1 MQALNITPEQFSLRRDHNLTREQFIALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNLGGAFICKM 120
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNLGGAFICKM 120
 QY 121 VPVQSTAVVTEILTMTCTIAVERHOGVLVHPFKMKQYTNRRAPFTMLGVVWLVAVIGSPM 180
 DB 121 VPVQSTAVVTEILTMTCTIAVERHOGVLVHPFKMKQYTNRRAPFTMLGVVWLVAVIGSPM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLEBWSVPVHOKIYTTFFILVILFLLPLMWMLILYSKIGYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICCLEBWSVPVHOKIYTTFFILVILFLLPLMWMLILYSKIGYE 240

QY 241 LMIKKRVGDSVLRTHGKEMSKIARKKKRAVIMVTVVAFVAVCWAPFHVHMMIEYSN 300
Db 241 LMIKKRVGDSVLRTHGKEMSKIARKKKRAVIMVTVVAFVAVCWAPFHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVIYAFMNFKNVLSAVCYCIVNKTSPAQ 360
Db 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVIYAFMNFKNVLSAVCYCIVNKTSPAQ 360
QY 361 RHGNSGITMWRKAKFSLRENPEETKGEAFSDGNIIEVKLCEOTEKKLKRHLALFRSE 420
Db 361 RHGNSGITMWRKAKFSLRENPEETKGEAFSDGNIIEVKLCEOTEKKLKRHLALFRSE 420
QY 421 LAENSPDLSGH 431
Db 421 LAENSPDLSGH 431
RESULT 3
AAB74773
ID AAB74773 standard; Protein; 431 AA.
XX AAB74773;
XX
DT 05-JUN-2001 (first entry)
DE Human G protein-coupled receptor protein AQ27 SEQ ID NO:1.
XX Human; G protein-coupled receptor; AQ27; nootropic; neuroprotective;
KW hypotensive; antirheumatic; antiallergic; cardiant; antianginal;
KW abortifacient; gene therapy; Alzheimer's disease; hypertension;
KW pregnancy termination; rheumatism; allergy; angina pectoris.
XX
OS Homo sapiens.
XX
XX WO200116316-A1.
XX
PD 08-MAR-2001.
XX
PF 24-AUG-2000; 2000WO-JP05684.
XX
PR 27-AUG-1999; 99JP-0241530.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
XX Watanabe T, Kikuchi K, Shintani Y;
XX
XX WPI; 2001-226689/23.
DR N-PSDB; AAF81818, AAF81819.
XX
XX Human brain-originated guanosine triphosphate protein-coupled receptor,
PT its salt and encoded gene, useful in (gene) diagnosis and development
PT of preventives and remedies for e.g. Alzheimer disease, hypertension
PT and rheumatism -
XX
PS Claim 1; Page 76-78; 84pp; Japanese.
XX
XX The present sequence represents a human G protein-coupled receptor
CC designated AQ27. AQ27 has nootropic, neuroprotective, hypotensive,
CC antirheumatic, antiallergic, cardiant, antianginal and abortifacient
CC activities. AQ27 can be used as a G protein-coupled receptor protein
CC agonist. G-protein-coupled receptor signal transducer and in gene
CC therapy. AQ27 can be sued in the diagnosis and development of
CC preventives and remedies for diseases associated with dysfunction of
CC AQ27 e.g. Alzheimer's disease, hypertension, pregnancy termination,
CC rheumatism, allergy and angina pectoris.
XX
SQ Sequence 431 AA;
Query Match 100.0%; Score 2235; DB 22; Length 431;
Best Local Similarity 100.0%; Pred. No. 2.3e-237;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPQFSRLLRDNLHNLTREQFIALYRLPLVVTPELPGRKALALVLTGLIFALAL 60
Db 1 MQALNITPQFSRLLRDNLHNLTREQFIALYRLPLVVTPELPGRKALALVLTGLIFALAL 60
QY 61 FGNALVYVVTSSKAMRTVTNIFICSALSDLLITPFCIPVTMLQNISDNWLGGAFCIKM 120
Db 61 FGNALVYVVTSSKAMRTVTNIFICSALSDLLITPFCIPVTMLQNISDNWLGGAFCIKM 120
QY 121 VFPVQSTAVVTILMTWCIAVERHOGLVHPFKMKQYTNRRAPTMLGVVWLVAVIVGSPM 180
Db 121 VFPVQSTAVVTILMTWCIAVERHOGLVHPFKMKQYTNRRAPTMLGVVWLVAVIVGSPM 180
QY 181 WHVQQLLEIKYDFLYEKEHICCLEEWTSPVHQIYITTFILVILFLPLMWMLILYSKIGYE 240
Db 181 WHVQQLLEIKYDFLYEKEHICCLEEWTSPVHQIYITTFILVILFLPLMWMLILYSKIGYE 240
QY 241 LMIKKRVGDSVLRTHGKEMSKIARKKKRAVIMVTVVAFVAVCWAPFHVHMMIEYSN 300
Db 241 LMIKKRVGDSVLRTHGKEMSKIARKKKRAVIMVTVVAFVAVCWAPFHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVIYAFMNFKNVLSAVCYCIVNKTSPAQ 360
Db 301 FEKEYDDVTIKMFAIVQIIGFNSICNPVIYAFMNFKNVLSAVCYCIVNKTSPAQ 360
QY 361 RHGNSGITMWRKAKFSLRENPEETKGEAFSDGNIIEVKLCEOTEKKLKRHLALFRSE 420
Db 361 RHGNSGITMWRKAKFSLRENPEETKGEAFSDGNIIEVKLCEOTEKKLKRHLALFRSE 420
QY 421 LAENSPDLSGH 431
Db 421 LAENSPDLSGH 431
RESULT 4
AAB48963
ID AAB48963 standard; Protein; 431 AA.
XX AAB48963;
XX
DT 27-MAR-2001 (first entry)
DE Human G protein-coupled receptor AXOR16.
XX
XX AXOR16; human; G protein-coupled receptor; 7TM receptor;
KW chromosome 11q12.2; infection; viral; bacterial; fungal;
KW protozoan; HIV-1; HIV-2; pain; cancer; diabetes; obesity;
KW anorexia; bulimia; osteoporosis; asthma; allergy; urinary retention;
KW acute heart failure; hypotension; hypertension; angina pectoris;
KW myocardial infarction; stroke; ulcer; migraine; vomiting;
KW psychotic disorder; neurological disorder; anxiety; schizophrenia;
KW manic depression; bipolar disorder; depression; delirium; dementia;
KW severe mental retardation; dyskinesia; Parkinson's disease;
KW Huntington's disease; Gilles de la Tourette's syndrome; gene therapy;
KW vaccine; drug screening; signal transduction; ds.
XX
OS Homo sapiens.
XX
XX WO2000078809-A1.
PN
XX 28-DEC-2000.
PD
XX 19-JUN-2000; 2000WO-US16869.
PF
XX 21-JUN-1999; 99US-0337105.
PR
XX (SMIK) SMITHKLINE BEECHAM CORP.
PA
XX Elshourbagy N, Shabon U;
PI
XX WPI; 2001-102699/11.
DR N-PSDB; AAC87890.
XX
PT New AXOR16, a G-protein coupled receptor polypeptide for screening

agonists and antagonists and for diagnosing and treating microbial infections, cancer, neurological disorders and asthma

Claim 1; Page 29-30; 31pp; English.

The invention relates to the human G protein-coupled receptor AXOR16 (AAB48963) to DNA encoding AXOR16 (AAC87690), and to AXOR16 fragments and variants. Like all G protein-coupled receptors, AXOR16 has 7 putative transmembrane domains and is involved in signal transduction. AXOR16 has homology with Gadus morhua neuropeptide (NPYRB) F and homology/ structural similarity with Danio rerio neuropeptide Y (NPYRYA). The human AXOR16 gene is located on chromosome 11q12.2. The invention also relates to expression vectors and host cells comprising AXOR16 DNA, to recombinant expression of AXOR16, and to an AXOR16-specific antibody. AXOR16 proteins and nucleotides may be used to treat a wide variety of disorders including bacterial, fungal, protozoal and viral infections, particularly HIV-1 or HIV-2 infections, pain; cancers; benign prostatic hypertrophy; diabetes; obesity; anorexia; bulimia; osteoporosis; asthma; allergies; urinary retention; acute heart failure; hypotension; hypertension; angina pectoris; myocardial infarction; stroke; ulcers; migraine; vomiting; psychotic and neurological disorders such as anxiety, schizophrenia, manic depression, depression, delirium, dementia, and severe mental retardation, and dyskinesias, such as Parkinson's disease, Huntington's disease or Gilles de la Tourette's syndrome. AXOR16 proteins, and nucleotides are useful as vaccines and AXOR16 proteins, nucleotides and antibodies may be used in screening compounds for their ability to modulate AXOR16 activity or expression. AXOR16 proteins are also useful for inducing an immunological response in a mammal against the above diseases, for antibody production, and to identify membrane bound or soluble receptors for AXOR16. AXOR16 nucleotides are also useful as diagnostic reagents and in chromosome localisation and tissue expression studies. The present sequence represents human AXOR16.

Sequence 431 AA;

Query Match 100.0%; Score 2235; DB 22; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITEQFSRLRDHNLTRQEQFIALYRLPLVYTFPELFGRAKLALVTGLIFALAL 60
 DB 1 MQALNITEQFSRLRDHNLTRQEQFIALYRLPLVYTFPELFGRAKLALVTGLIFALAL 60

QY 61 FGNALVFVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSDNWLGAFICKM 120
 DB 61 FGNALVFVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSDNWLGAFICKM 120

QY 121 VPFVQSTAVVTBILTMTCIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVLMVAVVGSFM 180
 DB 121 VPFVQSTAVVTBILTMTCIAVERHOGVLVHPFKMKQVYNNRAFTMLGVVLMVAVVGSFM 180

QY 181 WHVQOLEIKYDFLYEKEHICLLEWTSFVHQKIYTFILVLFLPLMWMLLYSKIGYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICLLEWTSFVHQKIYTFILVLFLPLMWMLLYSKIGYE 240

QY 241 LMKIKKVGDSVLRTHGKMSKIARKKRAVIMVMTVVALFACVAPPHVHVMISYN 300
 DB 241 LMKIKKVGDSVLRTHGKMSKIARKKRAVIMVMTVVALFACVAPPHVHVMISYN 300

QY 301 FEKEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENFKNVLSAVCYCVNKTFSPAQ 360
 DB 301 FEKEYDDVTIKMIFAIQVIGFSNSICNPVYAFNENFKNVLSAVCYCVNKTFSPAQ 360

QY 361 RHNSGITMWRKAKFSLRENPEVETKGBAPSDGNIKVLCEQTEKKLKRHLALFRSE 420
 DB 361 RHNSGITMWRKAKFSLRENPEVETKGBAPSDGNIKVLCEQTEKKLKRHLALFRSE 420

QY 421 LAENSPDLSGH 431
 DB 421 LAENSPDLSGH 431

RESULT 5
 AAU97735
 ID AAU97735 standard; Protein; 431 AA.
 XX
 AC AAU97735;
 XX
 DT 27-AUG-2002 (first entry)
 XX
 DE Human TGR346 polypeptide.
 XX
 KW Human; TGR346; receptor; G-protein coupled receptor; GPCR; TGR;
 KW TGR-associated disorder; signal transduction; renal failure; nephritis;
 KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
 KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;
 KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;
 KW circadian rhythm.
 XX
 OS Homo sapiens.
 XX
 PN WO200242458-A2.
 XX
 PD 30-MAY-2002.
 XX
 PF 21-NOV-2001; 2001WO-US43404.
 XX
 PR 22-NOV-2000; 2000US-252841P.
 PR 22-DEC-2000; 2000US-257636P.
 PR 12-JAN-2001; 2001US-261377P.
 PR 28-MAR-2001; 2001US-279554P.
 PR 29-MAR-2001; 2001US-280696P.
 XX
 PA (TULA-) TULARIK INC.
 XX
 PI Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;
 XX
 DR WPI; 2002-463633/49.
 DR N-PSDB; ABK86287.
 XX
 PT New isolated G-protein coupled receptor polypeptide, termed TGR, for
 PT diagnosis and treatment of diseases such as renal failure, nephritis,
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and
 PT sleep -
 XX
 PS Claim 33; Page 74; 98pp; English.
 XX
 CC The invention relates to a G-protein coupled receptor polypeptide (GPCR),
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of mutations associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification of
 CC modulators of GPCR signal transduction. This sequence represents the
 CC human TGR346 polypeptide.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2235; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;
 Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITEQFSRLRDHNLTRQEQFIALYRLPLVYTFPELFGRAKLALVTGLIFALAL 60
 DB 1 MQALNITEQFSRLRDHNLTRQEQFIALYRLPLVYTFPELFGRAKLALVTGLIFALAL 60

QY 61 FGNALVFVYVTRSKAMRTVTNIFICSLALSLLITFFCIPVTMLQNSDNWLGAFICKM 120

Db 61 FGNALVFYVTRSKAMRTVTNIFICSIALSDLLITPFCIPVTMLQNLSDNWLGAFICK 120
 QY 121 VPFVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRRAFTMLGVVWLVAIVGSPM 180
 Db 121 VPFVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRRAFTMLGVVWLVAIVGSPM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLSEWTSVPHQKIYTFILVILFLLPLVWMLLYSKIGYE 240
 Db 181 WHVQOLEIKYDFLYEKEHICCLSEWTSVPHQKIYTFILVILFLLPLVWMLLYSKIGYE 240
 QY 241 LWIKKRVGDSGLRTIHGKEMSKIAKKRAVIMMVTVVALFVAVCWAPFVHVHMIYEYN 300
 Db 241 LWIKKRVGDSGLRTIHGKEMSKIAKKRAVIMMVTVVALFVAVCWAPFVHVHMIYEYN 300
 QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFMNFENFKKNVLSAVCYCIVNKTSPQA 360
 Db 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFMNFENFKKNVLSAVCYCIVNKTSPQA 360
 QY 361 RHGNSGITMRKKAFFSLRENPEVETKGEAFSDGNIEVKLCEQTEKKLKRHLALFRSE 420
 Db 361 RHGNSGITMRKKAFFSLRENPEVETKGEAFSDGNIEVKLCEQTEKKLKRHLALFRSE 420
 QY 421 LAENSPDLSGH 431
 Db 421 LAENSPDLSGH 431

RESULT 6

AAU76035
 ID AAU76035 standard; Protein; 431 AA.

XX AC AAU76035;
 XX DT 08-MAY-2002 (first entry)

XX DE Human neuropeptide Y-like GPCR (NPY-like GPCR) protein sequence.
 KW Human; neuropeptide Y-like G protein-coupled; receptor; antibody;
 KW NPY-like GPCR; anorectic; antidiabetic; tranquilizer; hypertensive;
 KW hypotensive; analgesic; cytotatic; antiaesthetic; osteopathic;
 KW cardiovascular; antitumor; neurotropic; anticonvulsant; antiallergic;
 KW neuroleptic; antidote; antiparkinsonian; neuroprotective; gene therapy;
 KW antidepressant; antiaesthetic; diuretic; chromosome 11q12.2.

XX OS Homo sapiens.

XX PN WO200204518-A2.

XX PD 17-JAN-2002.

XX PF 06-JUL-2001; 2001WO-US21276.

XX PR 06-JUL-2000; 2000US-216523P.

XX PA (FARB) BAYER CORP.

XX PI Bloomquist BT, Zhelein L;

XX DR WPI; 2002-148166/19.

XX DR N-PSDB; ABK14956.

XX PT New neuropeptide Y-like G protein-coupled receptor, for detecting
 PT regulators for preventing, ameliorating, or correcting e.g. obesity,
 PT anorexia, diabetes, hypotension, hypertension, pain, cancers, or
 PT neurological disorders

XX PS Claim 10; Fig 1; 9lpp; English.

XX CC The present invention relates to a new polynucleotide encoding a
 CC polypeptide neuropeptide Y-like G protein-coupled receptor (NPY-like
 CC GPCR), comprising a 431 amino acid sequence, given in the specification.
 CC The polynucleotide of the invention is used to produce a NPY-like GPCR
 CC polypeptide. The complement of the polynucleotide can be used to detect

CC a coding sequence for a NPY-like GPCR. A reagent, preferably an antibody,
 CC that binds to NPY-like GPCR can be used to detect it. The NPY-like GPCR
 CC can be used to screen for agents that can regulate the activity of it. A
 CC reagent that binds to or regulates NPY-like GPCR can be used to reduce
 CC the biological activity of NPY-like GPCR, or treat a disorder associated
 CC with abnormal expression of NPY-like GPCR. The reagent is useful for
 CC preventing, ameliorating or correcting dysfunctions or diseases such as
 CC obesity, anorexia, bulimia, diabetes, hypotension, hypertension, cocaine
 CC withdrawal, pain, cancers, asthma, urinary retention, osteoporosis,
 CC angina pectoris, myocardial infarction, ulcers, allergies, congestive
 CC heart failure, memory enhancement, cardiac and cerebral vasospasm,
 CC pheochromocytoma, ganglioneuroblastoma, Huntington's disease, severe
 CC mental retardation, Alzheimer's disease, Parkinson's disease, psychotic
 CC and neurological disorders including anxiety, schizophrenia, manic
 CC depression, delirium, dementia and dyskinesias. The polypeptide may also
 CC be used in raising specific antibodies. The present amino acid sequence
 CC represents human NPY-like GPCR protein of the invention. This protein is
 CC encoded by the human NPY-like GPCR gene located on chromosome 11q12.2.

XX SQ Sequence 431 AA;

Query Match 100.0%; Score 2235; DB 23; Length 431;
 Best Local Similarity 100.0%; Pred. NO. 2.3e-237; Mismatches 0; Gaps 0;
 Matches 431; Conservative 0; Indels 0;

QY 1 MQALNITPEQFSRLLRDHNLTREQFIALYRLRPVLYTPELFGRAKLALVLTGVLFALAL 60
 Db 1 MQALNITPEQFSRLLRDHNLTREQFIALYRLRPVLYTPELFGRAKLALVLTGVLFALAL 60

QY 61 FGNALVFYVTRSKAMRTVTNIFICSIALSDLLITPFCIPVTMLQNLSDNWLGAFICKM 120
 Db 61 FGNALVFYVTRSKAMRTVTNIFICSIALSDLLITPFCIPVTMLQNLSDNWLGAFICKM 120

QY 121 VPFVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRRAFTMLGVVWLVAIVGSPM 180
 Db 121 VPFVQSTAVVTEILTMTCTIAVERHQGLVHPFKMKQVYTNRRRAFTMLGVVWLVAIVGSPM 180

QY 181 WHVQOLEIKYDFLYEKEHICCLSEWTSVPHQKIYTFILVILFLLPLVWMLLYSKIGYE 240
 Db 181 WHVQOLEIKYDFLYEKEHICCLSEWTSVPHQKIYTFILVILFLLPLVWMLLYSKIGYE 240

QY 241 LWIKKRVGDSGLRTIHGKEMSKIAKKRAVIMMVTVVALFVAVCWAPFVHVHMIYEYN 300
 Db 241 LWIKKRVGDSGLRTIHGKEMSKIAKKRAVIMMVTVVALFVAVCWAPFVHVHMIYEYN 300

QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFMNFENFKKNVLSAVCYCIVNKTSPQA 360
 Db 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFMNFENFKKNVLSAVCYCIVNKTSPQA 360

QY 361 RHGNSGITMRKKAFFSLRENPEVETKGEAFSDGNIEVKLCEQTEKKLKRHLALFRSE 420
 Db 361 RHGNSGITMRKKAFFSLRENPEVETKGEAFSDGNIEVKLCEQTEKKLKRHLALFRSE 420

QY 421 LAENSPDLSGH 431
 Db 421 LAENSPDLSGH 431

RESULT 7

ABG73503
 ID ABG73503 standard; Protein; 431 AA.

XX AC ABG73503;

XX DT 14-FEB-2003 (first entry)

XX DE Human G protein coupled receptor HGPBMY2 SEQ ID 17.

XX KW Human; G-protein coupled receptor; HGPBMY1; HGPBMY2; immunosuppressive;
 KW cardiant; neuroprotective; antiinflammatory; cytostatic; vulnery;
 KW vaccine; gene therapy; autoimmune; cardiovascular; neural; reproductive;
 KW haematopoietic; pulmonary; gastrointestinal; proliferation; cell cycle;
 KW birth defect; aberrant phosphorylation; acute phase response; receptor;

KW signal transduction; hyperimmune activity; inflammatory; hypercongenital;
 KW necrotic lesion; wound; organ transplant rejection.
 OS Homo sapiens.
 XX
 FN WO200268591-A2.
 PD 06-SEP-2002.
 XX
 XX 22-FEB-2002; 2002WO-US05281.
 XX
 XX 23-FEB-2001; 2001US-270792P.
 XX 23-FEB-2001; 2001US-270793P.
 PR 06-JUN-2001; 2001US-296427P.
 XX
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Feder J, Ramanathan C, Nelson T, Mintier G, Cacace A, Barber L;
 PI Kornacker M, Bol D;
 XX
 XX WPI; 2003-058304/05.
 DR N-PSDB; ABS57631.
 XX
 XX New human HGPBMY1 or HGPBMY2 polynucleotide and polypeptide, useful
 PT preventing, treating or ameliorating a disorder e.g., wound,
 PT cardiovascular disorder or transplant rejection -
 XX
 XX Claim 5; Figure 7; 316pp; English.
 XX
 CC This invention describes the novel human G-protein coupled receptors
 CC (GPCR's), HGPBMY1 or HGPBMY2 which have immunosuppressive, cardiant,
 CC neuroprotective, antiinflammatory, cytostatic and vulnerary activity
 CC and can be used in vaccines or for gene therapy. Pharmaceutical
 CC compositions comprising HGPBMY1 or HGPBMY2 polypeptides or their
 CC agonists or antagonists or modulators, or a HGPBMY1- or
 CC HGPBMY2-specific antibody are useful for preventing, treating or
 CC ameliorating a medical condition comprising autoimmune, cardiovascular,
 CC neural, reproductive, haematopoietic, pulmonary, gastrointestinal or
 CC proliferating disorder, a cell cycle or birth defect, a disorder related
 CC to aberrant phosphorylation, acute phase responses or signal transduction
 CC or to hyperimmune activity, an inflammatory or hypercongenital condition,
 CC a necrotic lesion, a wound, organ transplant rejection or a condition
 CC related to organ transplant rejection. This sequence represents the human
 CC HGPBMY2 protein described in the disclosure of the invention.
 XX
 SQ Sequence 431 AA;
 Query Match 100.0%; Score 2235; DB 24; Length 431;
 Best Local Similarity 100.0%; Pred. No. 2.3e-237; Indels 0; Gaps 0;
 Matches 431; Conservative 0; Mismatches 0;
 QY 1 MQALNITPEQSRLLRDHNLTRQFIYALYRLPLVYTPPELGRKALVLTGVLIFAL 60
 DB 1 MQALNITPEQSRLLRDHNLTRQFIYALYRLPLVYTPPELGRKALVLTGVLIFAL 60
 QY 61 FGNALVYVYVTRSKAMRTVNIIFCSIALSDLLITFCIPVTMLQNSDNLWGAFICKM 120
 DB 61 FGNALVYVYVTRSKAMRTVNIIFCSIALSDLLITFCIPVTMLQNSDNLWGAFICKM 120
 QY 121 VPVQSTAVVTEILTMTCAVERHQGLVHPFKMQYTNRRAPFTMLGVVWLVAIVGSPM 180
 DB 121 VPVQSTAVVTEILTMTCAVERHQGLVHPFKMQYTNRRAPFTMLGVVWLVAIVGSPM 180
 QY 181 WHVQOLEIKYDFLEKEHI CCLBEWSPVHOKIYTTFFILVFLPLMWMILHLSKIGYE 240
 DB 181 WHVQOLEIKYDFLEKEHI CCLBEWSPVHOKIYTTFFILVFLPLMWMILHLSKIGYE 240
 QY 241 LWIKRVGDSGLVLTTHGKMSKIARKKGAIVIMVTVVALFAVCWAPPHVVMIEYSN 300
 DB 241 LWIKRVGDSGLVLTTHGKMSKIARKKGAIVIMVTVVALFAVCWAPPHVVMIEYSN 300
 QY 301 FEKEYDDVTIKMIFAIQIIGPNSICNPITVAFMNFKNVLSAVCYCIYVNTKTSFPAQ 360

Db 301 FEKEYDDVTIKMIFAIQIIGPNSICNPITVAFMNFKNVLSAVCYCIYVNTKTSFPAQ 360
 QY 361 RHGNSGITMRRKAKFSLRENPFVETKGEAFSDGNIYVLCQTEQTEKKLKRHLALFRSE 420
 Db 361 RHGNSGITMRRKAKFSLRENPFVETKGEAFSDGNIYVLCQTEQTEKKLKRHLALFRSE 420
 QY 421 LAENSPDLSGH 431
 Db 421 LAENSPDLSGH 431
 RESULT 8
 AAY71309
 ID AAY71309 standard; Protein; 431 AA.
 XX
 AC AAY71309;
 XX
 DT 02-NOV-2000 (first entry)
 XX
 DE Human orphan G protein-coupled receptor hrUP4.
 XX
 KW Human; orphan G protein-coupled receptor; GPCR; hrUP4; drug screening;
 KW transmembrane receptor; expressed sequence tag; EST; signal cascade.
 XX
 OS Homo sapiens.
 XX
 FN WO200031258-A2.
 XX
 PD 02-JUN-2000.
 XX
 XX 13-OCT-1999; 99WO-US23687.
 XX 20-NOV-1998; 98US-0109213.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0121852.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123949.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 29-JUN-1999; 99US-0141448.
 PR 29-SEP-1999; 99US-0156555.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156634.
 PR 29-SEP-1999; 99US-0156653.
 PR 01-OCT-1999; 99US-0157280.
 PR 01-OCT-1999; 99US-0157281.
 PR 01-OCT-1999; 99US-0157282.
 PR 01-OCT-1999; 99US-0157293.
 PR 01-OCT-1999; 99US-0157294.
 PR 12-OCT-1999; 99US-0416760.
 PR 12-OCT-1999; 99US-0417044.
 XX
 XX (AREN-) ARENA PHARM INC.
 XX
 XX Chen R, Dang HT, Liaw CW, Lin I;
 XX
 DR WPI; 2000-400068/34.
 DR N-PSDB; AAD01136.
 XX
 PT Novel human orphan G protein-coupled receptors and the encoding cDNAs
 PT for use in the identification of G protein-coupled receptor agonists -
 XX
 PS Claim 74; Page 89-91; 102pp; English.
 XX
 CC The present amino acid sequence is the hrUP4, an endogenous human
 CC orphan G protein-coupled receptor (GPCR). The full length hrUP4 cDNA was
 CC cloned by RT-PCR with human brain cDNA as template. The hrUP4 PCR
 CC fragment obtained was an alternatively spliced form of the EST (expressed
 CC sequence tag) clone AI307658. The orphan GPCR of the invention, like

CC all GPCRs has seven transmembrane alpha helices with an extracellular
 CC N-terminus and an intracellular C-terminus. However, no endogenous
 CC ligands has yet been identified for the proteins of the invention. The
 CC orphan GPCRs may be used in the identification of their endogenous
 CC ligands, and to screen potential GPCR agonists and antagonists for use as
 CC pharmaceutical agents. The proteins may also be used in the study of
 CC GPCR-mediated signalling cascades, and to elucidate their precise role in
 CC normal and diseased human conditions. Nucleic acid encoding human orphan
 CC GPCRs may be used for tissue localisation expression analysis to provide
 CC information about their function in healthy and pathological states.

XX Sequence 431 AA;

Query Match 99.9%; Score 2232; DB 21; Length 431;
 Best Local Similarity 99.8%; Pred. No. 4.9e-237;
 Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQSRLLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLIFFALAL 60
 DB 1 MQALNITPEQSRLLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLIFFALAL 60
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
 QY 121 VPFVQSTAVVTEILTMTICIAVERHOGVLVHPFKMKQVYTNERRAFTMLGVWLVAIVGSPM 180
 DB 121 VPFVQSTAVVTEILTMTICIAVERHOGVLVHPFKMKQVYTNERRAFTMLGVWLVAIVGSPM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLEEWTSFVHQKIYTTFFILVILFLLPLMVMILLYSKIYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICCLEEWTSFVHQKIYTTFFILVILFLLPLMVMILLYSKIYE 240
 QY 241 LWIKKRVGDSGLRTINGKEMSKIARKKRAVINMVTVALEAVCWAPFHVHMMIEYSN 300
 DB 241 LWIKKRVGDSGLRTINGKEMSKIARKKRAVINMVTVALEAVCWAPFHVHMMIEYSN 300
 QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 DB 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 QY 361 RHGNSGITMRKKAFLRENPUETKGEAFSDGNIEVKLCQTEKKKLKRLHALFRSE 420
 DB 361 RHGNSGITMRKKAFLRENPUETKGEAFSDGNIEVKLCQTEKKKLKRLHALFRSE 420
 QY 421 LAENSPDLSGH 431
 DB 421 LAENSPDLSGH 431

RESULT 9

AAB02843
 ID AAB02843 standard; Protein; 431 AA.

AC AAB02843;

DT 22-AUG-2000 (first entry)

XX Human G protein coupled receptor hRUP4 protein SEQ ID NO:40.

DE Human; G protein coupled receptor; GPCR; transmembrane receptor;
 KW identification; agonist; screening; therapeutic; pharmaceutical;
 XX mutant.

OS Homo sapiens.

XX WO2000022131-A2.

PN 20-APR-2000.

XX 13-OCT-1999; 99WO-US24065.

PF 13-OCT-1998; 98US-0170496.

XX 13-OCT-1998; 98US-0170496.

PR 12-NOV-1998; 98US-0108029.
 PR 20-NOV-1998; 98US-0109213.
 PR 27-NOV-1998; 98US-0110060.
 PR 16-FEB-1999; 99US-0120416.
 PR 26-FEB-1999; 99US-0121852.
 PR 12-MAR-1999; 99US-0123944.
 PR 12-MAR-1999; 99US-0123945.
 PR 12-MAR-1999; 99US-0123946.
 PR 12-MAR-1999; 99US-0123948.
 PR 12-MAR-1999; 99US-0123949.
 PR 12-MAR-1999; 99US-0123951.
 PR 28-MAY-1999; 99US-0136436.
 PR 28-MAY-1999; 99US-0136437.
 PR 28-MAY-1999; 99US-0136439.
 PR 28-MAY-1999; 99US-0137127.
 PR 28-MAY-1999; 99US-0137131.
 PR 30-JUN-1999; 99US-0137567.
 PR 27-AUG-1999; 99US-0141448.
 PR 03-SEP-1999; 99US-0151114.
 PR 29-SEP-1999; 99US-0152524.
 PR 29-SEP-1999; 99US-0156633.
 PR 29-SEP-1999; 99US-0156655.
 PR 29-SEP-1999; 99US-0156634.
 XX (AREN-) ARENA PHARM INC.
 PA Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT,
 PI Gore M, Liaw CW, Lin I, Lowitz K, White C;
 XX WPI; 2000-317986/27.
 DR N-PSDB; AAA46037.
 XX Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 1; Page 119-120; 187pp; English.
 PS The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptor agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.
 XX Sequence 431 AA;

Query Match 99.9%; Score 2232; DB 21; Length 431;
 Best Local Similarity 99.8%; Pred. No. 4.9e-237;
 Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQALNITPEQSRLLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLIFFALAL 60
 DB 1 MQALNITPEQSRLLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLIFFALAL 60
 QY 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
 DB 61 FGNALVFYVYVTRSKAMRTVTNIFICSLSALSDLLITFFCIPVTMLQNSDNWLGGAFCIKM 120
 QY 121 VPFVQSTAVVTEILTMTICIAVERHOGVLVHPFKMKQVYTNERRAFTMLGVWLVAIVGSPM 180
 DB 121 VPFVQSTAVVTEILTMTICIAVERHOGVLVHPFKMKQVYTNERRAFTMLGVWLVAIVGSPM 180
 QY 181 WHVQOLEIKYDFLYEKEHICCLEEWTSFVHQKIYTTFFILVILFLLPLMVMILLYSKIYE 240
 DB 181 WHVQOLEIKYDFLYEKEHICCLEEWTSFVHQKIYTTFFILVILFLLPLMVMILLYSKIYE 240
 QY 241 LWIKKRVGDSGLRTINGKEMSKIARKKRAVINMVTVALEAVCWAPFHVHMMIEYSN 300
 DB 241 LWIKKRVGDSGLRTINGKEMSKIARKKRAVINMVTVALEAVCWAPFHVHMMIEYSN 300

QY 301 FEKEYDDVTIKMFAIVQIIIGFNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 FEKEYDDVTIKMFAIVQIIIGFNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 RHGNSGITMWRKAKFSLRENPFVETKGEAFSDGNIIEVKLCEOTEKPKLKRHLALFRSE 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 RHGNSGITMWRKAKFSLRENPFVETKGEAFSDGNIIEVKLCEOTEKPKLKRHLALFRSE 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 LAENSPDLSGH 431
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 LAENSPDLSGH 431
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 10

AA02853
 ID AAB02853 standard; Protein; 431 AA.

XX AC AAB02853;

DT 22-AUG-2000 (first entry)

DE Human G protein coupled receptor hRUP4 (V272K) protein SEQ ID NO:128.

KW Human; G protein coupled receptor; GPCR; transmembrane receptor;
 identification; agonist; screening; therapeutic; pharmaceutical;
 mutant.

XX OS Homo sapiens.

OS Synthetic.

XX WO2000022131-A2.

XX 20-APR-2000.

PF 13-OCT-1999; 99WO-US24065.

XX 13-OCT-1998; 98US-0170496.

PR 12-NOV-1998; 98US-0108029.

PR 20-NOV-1998; 98US-0109213.

PR 27-NOV-1998; 98US-0110060.

PR 16-FEB-1999; 99US-0120416.

PR 26-FEB-1999; 99US-0121852.

PR 12-MAR-1999; 99US-0123944.

PR 12-MAR-1999; 99US-0123946.

PR 12-MAR-1999; 99US-0123948.

PR 12-MAR-1999; 99US-0123949.

PR 28-MAY-1999; 99US-0136436.

PR 28-MAY-1999; 99US-0136437.

PR 28-MAY-1999; 99US-0136439.

PR 28-MAY-1999; 99US-0137127.

PR 28-MAY-1999; 99US-0137131.

PR 30-JUN-1999; 99US-0137567.

PR 27-AUG-1999; 99US-0151114.

PR 03-SEP-1999; 99US-0152524.

PR 28-SEP-1999; 99US-0152633.

PR 28-SEP-1999; 99US-0152633.

PR 29-SEP-1999; 99US-0156634.

XX (AREN-) ARENA PHARM INC.

XX Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;

PI Gore M, Liaw CW, Lin I, Lowitz K, White C;

XX WPI; 2000-317996/27.

XX N-PSDB; ARA46115.

XX Non-ergogenic, human G protein-coupled receptors for screening

PT receptor, inverse or partial agonists useful as therapeutic agents -

XX Example 2; Page 164-166; 187pp; English.

XX The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents,
 CC AAR46017 to AAR46126 and AAR02825 to AAR02829 represent sequences used in
 CC the exemplification of the present invention.

XX SQ Sequence 431 AA;

Query Match

Best Local Similarity 99.6%; Score 2226; DB 21; Length 431;

Matches 429; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MQALNITPEQFSRLDRDHLTREQFIYALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MQALNITPEQFSRLDRDHLTREQFIYALYRLPLVYTPPELPGRAKALVLTGVLIFALAL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 61 FGNALVFYVTRSKAMRTVNTNIFICSLSDLLITFFCIPVTMLQNI SDNMLGGAFICKM 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 FGNALVFYVTRSKAMRTVNTNIFICSLSDLLITFFCIPVTMLQNI SDNMLGGAFICKM 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 121 VPFVOSTAVVTEILTWTCTIAVERHQGLVHPFKKQYTNRRATMLGVVWLVAVIVGSPM 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 VPFVOSTAVVTEILTWTCTIAVERHQGLVHPFKKQYTNRRATMLGVVWLVAVIVGSPM 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYITFTILVILFLPLVMMLLYSKIGYE 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 WHVQOLEIKYDFLYEKEHICCLEEWTSPVHQKIYITFTILVILFLPLVMMLLYSKIGYE 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 241 LWIKKRVGDSVLTTHGKEMSKIAKKYRAVIMVTVVAVFVAVWPHVMMIEYSN 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 LWIKKRVGDSVLTTHGKEMSKIAKKYRAVIMVTVVAVFVAVWPHVMMIEYSN 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 301 FEKEYDDVTIKMFAIVQIIIGFNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 FEKEYDDVTIKMFAIVQIIIGFNSICNPIVYAFMNFKNVLSAVCYCIVNKTSPAQ 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 361 RHGNSGITMWRKAKFSLRENPFVETKGEAFSDGNIIEVKLCEOTEKPKLKRHLALFRSE 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 RHGNSGITMWRKAKFSLRENPFVETKGEAFSDGNIIEVKLCEOTEKPKLKRHLALFRSE 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 QY 421 LAENSPDLSGH 431
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 LAENSPDLSGH 431
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 11

AAU97740

ID AAU97740 standard; Protein; 433 AA.

XX AC AAU97740;

DT 27-AUG-2002 (first entry)

XX Mouse TGR346a polypeptide.

XX Mouse; TGR346a; receptor; G-protein coupled receptor; GPCR; TGR;

KW TGR-associated disorder; signal transduction; renal failure; nephritis;

KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;

KW diabetes insipidus; hyperprolactinemia; thirst disturbance; appetite;

KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;

XX circadian rhythm.

XX Mus sp.

XX WO200242458-A2.

XX 30-MAY-2002.

DB		421	LSENSTFGSGH	431	
		: :	:		
	RESULT 12				
	AAU97741				
ID	AAU97741	standard; Protein; 416	AA.		
XX	XX	AAU97741;			
XX	XX	27-AUG-2002	(first entry)		
DT	DT	Mouse	TGR346b polypeptide.		
XX	XX	DE	Mouse	TGR346b polypeptide.	
XX	XX	Mouse;	TGR346b; receptor; G-protein coupled receptor; GPCR; TGR;		
XX	XX	TGR-associated disorder; signal transduction; renal failure; nephritis;			
KW	KW	hyperthyroidism; hypogonadism; retinitis pigmentosa; growth disorder;			
KW	KW	diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;			
KW	KW	sleep disturbance; temperature regulation; blood pressure; hypothalamus;			
KW	KW	circadian rhythm.			
XX	XX	XX			
XX	XX	Mus	sp.		
XX	XX	WC200242458-A2.			
PN	PN	30-MAY-2002.			
XX	XX	21-NOV-2001;	2001WO-US43404.		
PF	PF	22-NOV-2000;	2000US-252841P.		
XX	XX	22-DEC-2000;	2000US-257636P.		
PR	PR	12-JAN-2001;	2001US-261377P.		
PR	PR	28-MAR-2001;	2001US-279554P.		
PR	PR	29-MAR-2001;	2001US-280696P.		
XX	XX	(TULA-)	TULARIX INC.		
PA	PA	Tian H,	Zhao J,	Chen J,	Cutler G,
XX	XX	An S,	Dai K,	Gupte JS;	
PI	PI				
XX	XX				

DR N-PSDE; ABR86293.
 PT New isolated G-protein coupled receptor polypeptide, termed TGR, for
 PT diagnosis and treatment of diseases such as renal failure, nephritis,
 PT hypothyroidism, diabetes insipidus, and disturbances of thirst and
 PT sleep -
 PS Claim 29; Page 79; 98pp; English.
 XX
 CC The invention relates to a G-protein coupled receptor polypeptide (GPCR)
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of mutations associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification
 CC of modulators of GPCR signal transduction. This sequence represents the
 CC mouse TGR345b polypeptide.
 XX
 XX Sequence 416 AA;
 Query Match 79.1%; Score 1767; DB 23; Length 416;
 Best Local Similarity 80.1%; Pred. No. 9.2e-186;
 Matches 339; Conservative 36; Mismatches 38; Indels 10; Gaps
 5 NITPEQFSALLRDHNLTRQQTALYRLRLVLTPELPGRKALVLVTGLIFALPFGNA 64

Db 4 NTABQLSALLRLHNLTRAQFIHAYGLRPLVLTLPQPARARLALLVGLMLFALALFGNA 63
QY 65 LVFVYVTRSKAMRTVNIIFCSIALSLLITFECPVTMLQISDNWLGAFICXVPTV 124
Db 64 LVVYVTRSKAMRTVNIIFCSIALSLLIVFPCIPVTMLQNVSDTMLGGAFICXVPTV 123
QY 125 QSTAVTEILTWTCTIAVERHQGLVHPFKMKQYTNRRFTMLGVVWLVAVIVGSPWHVQ 184
Db 124 QCTAIVTEILTWTCTIAVERHQGLVHPFKMKQYTNQRAFTMLGVVWLVAVIIGSPWHVQ 183
QY 185 QLEIKYDFLYEKEHCCLLEWTSVHQKIYTTFTILVILFLLPMWMLIYSGIYELWTK 244
Db 184 RLEIKYDFLYEKEHCCLLEWTSVHQKIYTTFTILVILFLLPMWMLIYSGIYELWTK 243
QY 245 KRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAFVAVCPVAVHVMIEYSNFEKE 304
Db 244 KRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAFVAVCPVAVHVMIEYSNFEKE 303
QY 305 YDVTIKMFAIVQIIGFSNSICNPVAFMNFKNVLSAVCYCIVNKTSPAQRHGN 364
Db 304 YDEVTKMFAIVQIIGFSNSICNPVAFMNFKNVLSAVCYCIVNKTSPAQRHGN 363
QY 365 SGITWMRKAKESLBNPVEKGFAPSDGNIKCEOTBEKKLKHHLALFRSELAEN 424
Db 364 SGITWMRKAKESLBNPVEKGFAPSDGNIKCEOTBEKKLKHHLALFRSELAEN 424
QY 425 SPL 427
Db 414 CPL 416

RESULT 13
ABP82011
ID ABP82011 standard; Protein; 245 AA.
AC ABP82011;
DT 04-MAR-2003 (first entry)
DE Human G protein-coupled receptor GPR103 protein SEQ ID NO:510.
KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
KW G protein-coupled receptor modulator; antibody; immune-related disease;
KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
KW immunological-related cell proliferative disease; autoimmune disease;
KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
KW ulcer.
OS Homo sapiens.
XX
XX WO200261087-A2.
FN
XX
XX PD 08-AUG-2002.
XX
PF 19-DEC-2001; 2001WO-US50107.
XX
PR 19-DEC-2000; 2000US-257144P.
XX
PA (LIFE-) LIFESPAN BIOSCIENCES INC.
XX
PI Burmer GC, Roush CL, Brown JP;
XX
XX WPI; 2003-046718/04.
DR N-PSDB; ABZ42861.
XX
XX New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,

PT cancer or autoimmune diseases -
XX Disclosure; Fig 1; 523pp; English.
CC The present invention describes antigenic peptides (I) comprising:
CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids. Also described: (1) an assay for the detection of a particular
CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity
CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC an antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention.
XX Sequence 245 AA;

Query Match 56.7%; Score 1268; DB 24; Length 245;
Best Local Similarity 99.2%; Pred. No. 4.7e-131;
Matches 243; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
QY 186 LEIKYDFLYEKEHCCLLEWTSVHQKIYTTFTILVILFLLPMWMLIYSGIYELWTK 245
Db 1 MKIKYDFLYEKEHCCLLEWTSVHQKIYTTFTILVILFLLPMWMLIYSGIYELWTK 60
QY 246 RVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAFVAVCPVAVHVMIEYSNFEKE 305
Db 61 RVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAFVAVCPVAVHVMIEYSNFEKE 120
QY 306 DDVTIKMFAIVQIIGFSNSICNPVAFMNFKNVLSAVCYCIVNKTSPAQRHGN 365
Db 121 DDVTIKMFAIVQIIGFSNSICNPVAFMNFKNVLSAVCYCIVNKTSPAQRHGN 180
QY 366 GITWMRKAKESLBNPVEETKGEAFSDGNIKCEOTBEKKLKHHLALFRSELAEN 425
Db 181 GITWMRKAKESLBNPVEETKGEAFSDGNIKCEOTBEKKLKHHLALFRSELAEN 240
QY 426 PLDSG 430
Db 241 PLDSG 245

RESULT 14
AAI79377
ID AAI79377 standard; Protein; 417 AA.
XX
XX AAI79377;
XX
DT 01-AUG-2000 (first entry)
XX
XX Rat neuropeptide FF (NPFF2) receptor.
XX
XX Neuropeptide FF receptor; NPFF2 receptor; rat; antiinflammatory;
XX antiasthmatic; antidiabetic; immunostimulant; immunosuppressive;
XX neurotropic; neuroprotective; analgesic; anorectic; antipsychotic;
XX antiaddictive; antimigraine; hypertensive; hypotensive; cardiant;
XX antiasthmatic; therapy; G protein coupled receptor.

OS Rattus norvegicus.
 XX Key Location/Qualifiers
 FT Domain 44..71
 FT Domain /note= "transmembrane domain I"
 FT Domain 81..104
 FT Domain /note= "transmembrane domain II"
 FT Domain 123..141
 FT Domain /note= "transmembrane domain III"
 FT Domain 161..180
 FT Domain /note= "transmembrane domain IV"
 FT Domain 220..243
 FT Domain /note= "transmembrane domain V"
 FT Domain 271..296
 FT Domain /note= "transmembrane domain VI"
 FT Domain 314..339
 FT Domain /note= "transmembrane domain VII"
 FT Modified-site 10
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 18
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 113
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 195
 FT Modified-site /note= "N-glycosylated"
 FT Modified-site 154
 FT Modified-site /note= "O-phosphorylated"
 FT Modified-site 263
 FT Modified-site /note= "O-phosphorylated"
 FT Modified-site 264
 FT Modified-site /note= "O-phosphorylated"
 XX
 PN W0200018438-A1.
 XX
 PD 06-APR-2000.
 XX
 PF 24-SEP-1999; 99WO-US22384.
 XX
 PR 25-SEP-1998; 98US-0161113.
 PR 22-FEB-1999; 99US-0255368.
 XX
 PA (SYNA-) SYNAPTIC PHARM CORP.
 XX
 PI Gerald CPG, Jones KA, Bonini JA, Borowsky B;
 XX
 DR WPI; 2000-293017/25.
 DR N-PSDB; AAZ94669.
 XX
 PT Nucleic acid encoding a mammalian neuropeptide FF (NPFF) receptor,
 PT useful for treatment of e.g pain, obesity, diabetes, hypertension,
 PT hypotension, hypoglycemia, respiratory disorders
 XX
 PS Claim 21; Fig 23A-B; 253pp; English.
 XX
 CC The present sequence is that of rat neuropeptide FF (NPFF2)
 CC receptor, as deduced from a cDNA clone (see AAZ94669). Isolated
 CC from rat spinal cord cDNA. High levels of rat NPFF2 mRNA are
 CC found in the central nervous system. Expression patterns suggest
 CC roles for NPFF2 in neuroendocrine regulation, and in regulation of
 CC circadian rhythm, regulation of appetite and other functions
 CC modulated by the hypothalamus. A possible role in regulation of
 CC cardiovascular function is also suggested. High levels in the
 CC amygdala suggest a role in modulation of mood, fear, phobia and
 CC anxiety, and NPFF2 may be a target for treatment of depression and
 CC other neuropsychiatric disorders. The invention provides rat
 CC and human NPFF polypeptides and polynucleotides, vectors, host
 CC cells, antibodies, nucleic acid probes, antisense oligonucleotides,
 CC transgenic animals, methods of isolating mammalian NPFF receptors,
 CC methods of treating an abnormality associated with NPFF receptor
 CC activity, methods of determining binding of compounds to NPFF
 CC receptors, methods of identifying agonists and antagonists of NPFF
 CC receptors, and the agonists and antagonists obtained. Claimed
 CC methods of treating an abnormality that is alleviated by

CC increasing/decreasing NPFF activity involve administering an NPFF
 CC receptor agonist/antagonist. The abnormality is a lower urinary
 CC tract disorder, an epinephrine release disorder, a gastrointestinal
 CC disorder, irritable bowel syndrome, a cardiovascular disorder, an
 CC electrolyte balance disorder, diuresis, hypertension, hypotension,
 CC diabetes, hypoglycemia, a respiratory disorder, asthma, a
 CC reproductive function disorder, an immune disorder, an endocrine
 CC disorder, a musculoskeletal disorder, a neuroendocrine disorder, a
 CC cognitive disorder, a memory disorder, a sensory modulation and
 CC transmission disorder, a motor coordination disorder, a sensory
 CC integration disorder, obesity, pain, psychotic behavior, a sensory
 CC morphine tolerance, nicotine addiction, opiate addiction,
 CC affective disorder or migraine (all claimed).
 XX
 SQ Sequence 417 AA;
 Query Match 22.7%; Score 508; DB 21; Length 417;
 Best Local Similarity 31.8%; Pred. NO. 6.9e-47;
 Matches 119; Conservative 72; Mismatches 145; Indels 38; Gaps 9;
 QY 54 LIFALALFGNALVYVVTFSKAMRTVTNIFICSLALSDLLITFFCIPVTMLQNISDNWLG 113
 DB 54 LIFFLCWGNTVVCVFVIRNRYHTVTNFFIFNLALSDLLVGFCEPITLLDNIAGWPF 113
 QY 114 GAFICKAVPFVQSTAVVTETILTMTCTIAVERHOGLVHPFKMKQYTNRRFTMLGVVWLVA 173
 DB 114 GSSMCKISGLVQGISVAASVFTLVAIVDRFCVYVFPFK--LTVKTFVMIIVWGLA 171
 QY 174 VIVGSP---MHWVQLEIKYDFLYEKH-----ICLEWTSPPVHKIYTTFTILVILFL 224
 DB 172 ITIMTFSAINLHVQ--EEKYVRVRLSSHNNKISTVYVWCRDWPNQEMRRIYTTVLFTATYL 229
 QY 225 LPLVMYLILYSKIGYELWIKRVGDSVLTTHGKEMSKIARKKCAVIMVTVVALFAV 284
 DB 230 APLSLIVINRYARIGASLFTSAHSGK-----QRLEQWVHVKKKQKVIKLLTVALLFIL 284
 QY 285 CWAPFHVHVHMIEYSNFEKEYDDVTIKMIFAIVQIIGFSNINPTIVYAFMNFKNVL 344
 DB 285 SWLPLWTLMLSDYADLSNKLRLVINIYVYPAHMLAFCSNVNPIYGFNFNFRSGFQ 344
 QY 345 SAVCYIVNKTSPAQRHNSGITWMRKAKPSLR-----ENPVEETKG-EAFS 392
 DB 345 DAFQFC--OKVKPQEBAYG-----LRKRNLIDINTSGLLVHEPASQNPSENLGCRKSA 396
 QY 393 DGNIEVKLCQETEE 406
 DB 397 DNPTQESLMEETGE 410
 RESULT 15
 AAW81358
 ID AAW81358 standard; Protein; 420 AA.
 XX
 AC AAW81358;
 XX
 DT 30-MAR-1999 (first entry)
 XX
 DE Human 7-transmembrane receptor HLNAR77.
 XX
 KW Human; transmembrane receptor; antagonist; infection; bacterium; fungus;
 KW protozoan; virus; HIV; pain; cancer; anorexia; bulimia; asthma; ulcer;
 KW Parkinson's disease; heart failure; hypotension; hypertension; asthma;
 KW urinary retention; osteoporosis; angina pectoris; myocardial infarction;
 KW allergy; benign prostatic hypertrophy; neurological disorder.
 XX
 OS Homo sapiens.
 XX
 PN EP884387-A2.
 XX
 PD 16-DEC-1998.
 XX
 XX 09-JUN-1998; 98EP-0304580.
 XX

```

PR 13-JAN-1998; 98US-0006140.
PR 11-JUN-1997; 97US-0049332.
PR 02-DEC-1997; 97US-0067253.
XX
XX
XX (SMIK ) SMITHLINE BEECHAM CORP.
XX
XX Elshourbagy N, Sathe G;
XX
XX WPI; 1999-026581/03.
XX N-PSDB; AAV68484.
XX
XX New DNA encoding 7-trans-membrane receptor polypeptide HLMAR77 -
XX used to treat, diagnose and prevent infections, pain, cancers,
XX anorexia, asthma, Parkinson's disease, acute heart failure,
XX osteoporosis, ulcers, allergies and psychotic disorders
XX
XX Claim 11; Page 8-9; 27pp; English.
XX
XX This sequence represents the human 7-transmembrane receptor HLMAR77.
XX The protein can be used to isolate agonists and antagonists. These can
XX be used as active agents in the treatment of infections (e.g. bacterial,
XX fungal, protozoal and viral infections, particularly HIV-1 or HIV-2),
XX pain, cancers, anorexia, bulimia, asthma, Parkinson's disease, acute
XX heart failure, hypotension, hypertension, urinary retention,
XX osteoporosis, angina pectoris, myocardial infarction, ulcers, asthma,
XX allergies, benign prostatic hypertrophy and psychotic and neurological
XX disorders.
XX
XX Sequence 420 AA;
XX
Query Match 22.5%; Score 502; DB 20; Length 420;
Best Local Similarity 30.1%; Pred. No. 3.2e-46;
Matches 128; Conservative 87; Mismatches 162; Indels 48; Gaps 12;
QY 14 LLRDHNLTRQFIALYRLRPLVYTPPLPGRAKLALVGLGVLPAALFGNALVFVVVTRS 73
DB 26 LYSINIT--YVNYLHQPV-----AAFIISYFLIFLCWNGNTVVCFIIVNRN 73
QY 74 KAMFTVNTIFCSLALSLLITFFCIPVTMLQNSDNLGGLGAFCKMVPFVQSTAVVTEI 133
DB 74 KMHFTVNTLNLALISDLLVGIFCMPTLLDNIAGWPGNTWCKISGLVQGISVAASV 133
QY 134 LTMTCIAVERHOGVLHPFKMKQYTNRRFTMLGVWLVAVIVGSP---MWHVQOLEIKY 190
DB 134 PTLVAIAVDRFCQVYVPPPKR--LTIKTAFVIMIIWVLAITIMSPSAVMLHVQE-EKYY 190
QY 191 DFLYEKEH-----ICCLEEWTSVHQKIYTTFILVFLPLMVMILLYSKIGYELWIKK 245
DB 191 RVRLNSQNKTSPTVYWCREDWPNQEMRKIYTVLVFANIVLAPLSLIVIMYGRIGISLP--- 247
QY 246 RVGGSVLRTIHGEMSK-----IARKKRAVIMVTVVVALFVAVCWAPPHVVHMMIYNSNF 301
DB 248 ---RAAVPHT--GRKNQEQHWVSRKKQKIIKMLLIALLFLLSWLPLWLTMLMSDYADL 302
QY 302 EKEVDDVTIKMIFAIVQIIIGFSNSICNPIVYAFVNEFKNVLSAVCYIVNKTFSPAQR 361
DB 303 SPNELQIINTIYIFFAHLAFGNSVNPDIYGFNENFRGRQFAFQLQLOCKRAKPMEA 362
QY 362 HGNSGITMMRKAKFSLFENPVEETKGEA-FSDGNIYVKLCQTEEEKKKLKHLLALFRSE 420
DB 363 YA-----LKAKSHVLINTSNQLVQESTFQNPHEGTLLYRKSAKPKQOE----LVMEER 410
QY 421 LAENS 425
DB 411 LKETT 415

```

Search completed: October 28, 2003, 10:12:10
Job time : 89 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 10:08:34 ; Search time 42 Seconds
(without alignments)
986.874 Million cell updates/sec

Title: US-10-070-24lb-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRLRDHNL.....RHLALFSELAEKNSFLDSGH 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	469.5	21.0	381	I39187	neuropeptide Y/pep
2	466.5	20.0	375	S63685	neuropeptide Y rec
3	426	19.1	382	S27388	neuropeptide Y rec
4	425.5	19.0	375	G02300	pancreatic polypep
5	425.5	18.0	375	I39182	neuropeptide Y/pep
6	425.5	19.0	449	A41738	neuropeptide Y rec
7	425	19.0	382	B46133	neuropeptide Y/pep
8	421.5	18.9	384	A45490	neuropeptide Y/pep
9	420.5	18.8	349	S12863	G protein-coupled
10	420.5	18.8	349	I59336	galanin receptor 1
11	412.5	18.5	366	S71152	neuropeptide Y/pep
12	402.5	18.0	427	S50150	gastric CCK-A rece
13	401	17.9	519	S17783	tachykinin recepto
14	400.5	17.9	423	B40470	glucocorticoid-ind
15	398	17.8	457	T29741	hypothetical prote
16	386.5	17.3	407	S23510	neurokinin 1 recep
17	384	17.2	402	I58595	neurokinin 2 recep
18	383.5	17.2	407	JQ1274	neurokinin 1 recep
19	382.5	17.1	444	A42685	cholecystokinin re
20	381.5	17.1	423	D40470	glucocorticoid-ind
21	379	17.0	428	JN0692	cholecystokinin ty
22	378.5	16.9	430	I51898	cholecystokinin A
23	378	16.9	385	S55224	neurokinin 3 recep
24	377.5	16.9	423	JC7677	allatostatin recep
25	377.5	16.9	584	JC7809	sulfakinin recepto
26	375.5	16.8	412	T22076	hypothetical prote
27	375.5	16.8	436	JC5599	cholecystokinin-A
28	372	16.6	452	A34916	neurokinin 3 recep
29	370.5	16.6	384	S20303	neurokinin 2 recep

30	370.5	16.6	399	2	A46632	bombesin-like pept
31	370.5	16.6	399	2	S29480	bombesin receptor
32	369	16.5	384	2	I57957	neurokinin 2 recep
33	367.5	16.4	407	2	A34357	neurokinin 1 recep
34	367.5	16.4	491	2	C40470	glucocorticoid-ind
35	366.5	16.4	407	2	S20304	neurokinin 1 recep
36	365	16.3	398	1	JQ1059	neurokinin 2 recep
37	364.5	16.3	384	2	I57682	bombesin/ GRP rece
38	362.5	16.2	384	1	S00516	neurokinin 2 recep
39	362.5	16.2	384	2	A39003	bombesin/gastrin-r
40	362.5	16.2	465	1	JQ1517	neurokinin 3 recep
41	361	16.2	370	1	I52315	G protein-coupled
42	360.5	16.1	390	2	B41007	bombesin-receptor
43	359.5	16.1	390	2	A36737	neurokinin 2 recep
44	359	16.1	384	2	A41007	gastrin-releasing
45	359	16.1	390	2	JH0374	bombesin receptor

ALIGNMENTS

RESULT 1

I39187

N/Alternate names: neuropeptide YY receptor Y2 - human

C/Species: Homo sapiens (man)

C/Date: 01-Mar-1996 #sequence revision 01-Mar-1996 #text_change 20-Apr-2000

C/Accession: I39187; I39163; G02301

J. Biol. Chem. 270, 26758-26761, 1995

A/Title: Expression cloning and pharmacological characterization of a human hippocampal

A/Reference number: I39187; MUID:96070760; PMID:7592910

A/Accession: I39187

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-381 <GR>

A/Cross-references: EMBL:U36269; NID:G1063633; PIDN:AA050281.1; PID:G1063634

R. Rose, P.M.; Fernandes, P.; Lynch, J.S.; Frazier, S.T.; Fisher, S.M.; Kodukula, K.; Kie

J. Biol. Chem. 270, 22661-22664, 1995

A/Title: Cloning and functional expression of a cDNA encoding a human type 2 neuropeptic

A/Reference number: I39183; MUID:96032678; PMID:7559383

A/Accession: I39163

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-133 'A', 135-381 <ROS>

A/Cross-references: EMBL:U32500; NID:G1000750; PIDN:AAA93170.1; PID:G1000751

R. Yan, H.; Yang, J.; Marasco, J.; Yamaguchi, K.; Brenner, S.; Collins, F.; Karbon, W.

submitted to the EMBL Data Library, December 1995

A/Reference number: H01019

A/Accession: G02301

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: mRNA

A/Residues: 1-171 'G', 173 'R', 175-201 'P', 203-208 'A', 210-381 <YAN>

A/Cross-references: EMBL:U42389; NID:G1314329; PIDN:AAB07760.1; PID:G1314330

C/Genetics:

A/Gene: GDB:NPY2R

A/Cross-references: GDB:4365607; OMIM:162642

A/Map position: 4q31-4q31

C/Superfamily: neurokinin 1 receptor

C/Keywords: appetite; G protein-coupled receptor; glycoprotein; lipoprotein; thiolester

F/49-76/Domain: transmembrane #status predicted <TM1>

F/87-113/Domain: transmembrane #status predicted <TM2>

F/166-186/Domain: transmembrane #status predicted <TM4>

F/221-237/Domain: transmembrane #status predicted <TM5>

F/269-291/Domain: transmembrane #status predicted <TM6>

F/305-328/Domain: transmembrane #status predicted <TM7>

F/123-203/Disulfide bonds: #status predicted

F/342/Binding site: palmitate (Cys) (covalent) #status predicted

F/372/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match

Best Local Similarity 21.0%; Score 469.5; DB 2; Length 381;

Matches 113; Conservative 75; Mismatches 123; Indels 53; Gaps 13;

QY 153 MKWQYNRAFTNMLGVVWLJAVIVGSPMHWVOQL-----EIKYDFLYEKHEICLEBWT 206
:
|
D**b** 141 RGRWRNNRHACGITVIWGFMAGCTPLMMYSVLTDPEFFKNISLDSYGK--YYCLEDDFP 198
:
|
Qy 207 SPVHOKIYTTFILVFLLPMLVMMLLSKYGLWIWKRVGDGSLVTIGHKEMSKTAR 266
:
|
Db 199 EDKPFLSVTTTLFILOYLGPLCFIFVCYTKI--FURLKER-----NNMMDKIRD 245
:
|
Qy 267 KK-----KRIVIMMTTVVALFAVCWAPPHVVHMMEIYSNFKEKEYDDDTIKMIFAIVOII 320
:
|
Db 246 NKYRSSETKRINIMLSLVGWFCALCWLFFFNLNVFDWNH--EAVATCNHLLELIChLT 303
:
|
Qy 321 GFSNSNICPIVYAfmENFKKVLSAvcYC 350
:
|
Db 304 AMISTCVNPIFYGFUNKPNQRDLQOFFFNPC 333
:

RESULT 12

S\$0150

gastric CCK-A receptor - rabbit

C;Species: Oryctolagus cuniculus (domestic rabbit)

C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 20-Apr-2000

R;Accession: S\$0150

E;Reuben, M.; Rising, L.; Prinz, C.; Hersey, S.; Sachs, G.

Biochim. Biophys. Acta 1219, 321-327, 1994

A;Title: Cloning and expression of the rabbit gastric CCK-A receptor.

A;Reference number: S\$0150; PMID:95002144; PMID:7918628

A;Accession: S\$0150

A>Status: preliminary

A:Molecule type: mRNA

A;Residues: 1-427 <REU>

C;Superfamily: neurokinin 1 receptor

Query Match 18.0%; Score 402.5; DB 2; Length 427;
Best Local Similarity 27.0%; Pred.No.5.5e-25;
Matches 95; Conservative 72; Mismatches 126; Indels 59; Gaps 7;

Qy 49 VLGVLIFALAFGNALFYVVRTSKAMRTVINIFICSLASLDLITFFCIPTVMIONIS 108
:
|
Db 45 ILLSYLFLLSvlgNTLViTLIRNKERTVINFILSLASLDLMCLFCMPNFLINFLL 104
:
|
Qy 109 DNWLGGAFICKWPFGVSSTAVTEILLTCIAVERHQGLVPHPKW-KWYNRRRAFTMLG 167
:
|
Db 105 KDFIFSALCKTTTTLGMTGSYSTLNVAISLERYGAI CKPLQSrvwQ-TKSHALKVIA 163
:
|
Qy 168 VWVLVAVIVGSPMHWOOLEIKYPDEYEKEHCICLEBTSPVHQKIYTTFILVFLLPL 227
:
|
Db 164 ATWCLSAIMTP-YPIYSNLVPFKTNQTANMCrFLLPSDMVMOQAHTFLLLjilFlPG 222
:
|
Qy 228 MWMLILysKTGYELW-----KKRW----- 247
:
|
Db 223 INVMVAYGISLEYIQGIKFEDASKKSACKERKAOSTSGRPEDNDGCYLRQSKPTRQLELQ 282
:
|
Qy 248 ---GDGsvlrrTHgkEMSkiARKKkaVImmTvAlPAVCwapPhVmhmIEYSnfEk 303
:
|
Db 283 QLsgggGGrVsRIHSssAAalMAKgVRInLMviVvlFcWmp-----IfSANAWR 335
:
|
Qy 304 EYDVtIk----MiFaivOIgfSNsicnPiVyafNmEnfKnvlsAvCYc 350
:
|
Db 336 AYDTVSAERRLSGTpIsFiLLsyTSscVNpiLycfNMKeFrIlGfMATPFCC 387
:
|

RESULT 13

\$17783

tachykinin receptor homolog DTKR - fruit fly (*Drosophila melanogaster*)

C;Species: *Drosophila melanogaster*

C;Date: 22-Nov-1996 #sequence_revision 22-Nov-1996 #text_change 20-Apr-2000

R;Li, X.J.; Wolfgang, W.; Wu, Y.N.; North, R.A.; Forte, M.
EMBO J. 10; 3221-3229, 1991

A;Title: Cloning, heterologous expression and developmental regulation of a *Drosophila* tachykinin receptor.

A;Reference number: \$17783; PMID:9200772; PMID:1717263

A;Accession: \$17783

Result No.	Query			ID	Description
	Score	Match	Length		
1	508	22.7	417	1	NFR2 RAT
2	592	22.5	522	1	NFR2 HUMAN
3	469.5	21.0	381	1	NYR2 HUMAN
4	468.5	21.0	381	1	NYR2 WACMU
5	467.5	20.9	382	1	NYR2 PIG
6	467	20.9	384	1	NYR2 BOVIN
7	466.5	20.9	385	1	NYR2 MOUSE
8	460.5	20.6	460	1	OXR2 MOUSE
9	459.5	20.6	460	1	OXR2 RAT
10	457	20.4	381	1	NYR2 CAVPO
11	455	20.4	444	1	OXR2 HUMAN
12	451	20.2	444	1	OXR2 CANFA
13	446.5	20.0	375	1	NYR2 MOUSE
14	446	20.0	371	1	NYR2 RABIT
15	440	19.7	385	1	NYR2 CHICK
16	437	19.6	432	1	NFL1 RAT
17	432.5	19.4	430	1	NFL1 HUMAN
18	431.5	19.3	346	1	GAUR RAT
19	431	19.3	383	1	NYR1 CAVPO
20	430.5	19.3	348	1	GAUR MOUSE
21	430.5	19.3	375	1	NYR1 RAT
22	429	19.2	371	1	NYR1 MOUSE
23	426	19.1	382	1	NYR1 MOUSE
24	425.5	19.0	375	1	NYR1 HUMAN
25	425.5	19.0	449	1	NYR1 DROME
26	425	19.0	382	1	NYR1 RAT
27	424	19.0	383	1	NYR1 PIG
28	423.5	18.9	382	1	NYR1 CANFA
29	422.5	18.9	425	1	OXR1 HUMAN
30	421.5	18.9	384	1	NYR1 HUMAN
31	420.5	18.8	349	1	GAUR HUMAN
32	412.5	18.5	366	1	NYR1 XENLA
33	412.5	18.5	416	1	OXR1 RAT

```

FT DOMAIN 67 82 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 83 103 2 (POTENTIAL).
FT DOMAIN 104 119 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 120 140 3 (POTENTIAL).
FT DOMAIN 141 160 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 161 181 4 (POTENTIAL).
FT DOMAIN 182 217 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 218 238 5 (POTENTIAL).
FT DOMAIN 239 274 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 275 295 6 (POTENTIAL).
FT DOMAIN 296 310 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 311 331 7 (POTENTIAL).
FT DOMAIN 332 417 CYTOPLASMIC (POTENTIAL).
FT DISULFID 118 206 BY SIMILARITY.
FT CARBOHYD 8 8 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 417 AA; 47710 MW; DFEB6AC63AF2AAD6 CRC64;

Query Match 22.7%; Score 508; DB 1; Length 417;
Best Local Similarity 31.8%; Pred. No. 7.5e-29;
Matches 119; Conservative 72; Mismatches 145; Indels 38; Gaps 9;

.QY 54 LIFALALFGNALVYVVTSSKAMRTVNIPIFICSLASDLITFFCIPVTMLQMSIDNWLG 113
DB 54 LIFELCWGNVTVCVIRNRYMHTVTFIFNLASDLLVGVFCMPITLLDNIAGWPF 113

.QY 114 GAFICMKVPVQSTAVTEILTWTCTIAVERHQGLVHPFKQVYNNRAFTMLGVVWLVA 173
DB 114 GSNMCKISGIVQGISVAASFTLVAVDFRCVYFPFK--LTVKTAFAVMIVVWGIA 171

.QY 174 VIVGSP---NMHWQOIEIKYDFLYEKH-----ICLEWTSFVHKIYTTTFFILVFL 224
DB 172 ITWTBSAHLVQ--BEKYVYVRLSHNKTSTVYWCREDWPNQENRITYTVLTIATYL 229

.QY 225 LPLMWMLLYSKYELWIKRKGVDGSLVTHGKEMSKIARKKRAVMVTVVVALPAV 284
DB 230 APLSLIVMYARIGASLFTKSAHSTGK-----ORLEQWHSVKKQKQVIMKLLTVALLFTL 284

.QY 285 CWAPFFHVHMIYSNPEKEYDDVTIKMFAIVQIIGFNSICNPVYAFPMENFKQV 344
DB 285 SWLPLTWLMLSDYADLSFNKLRVINYVYPPAHWLAFCNSSVNPVLYGFFNFRSGPQ 344

.QY 345 SAVCYCIKNTFPQARHGNSGIMTKKAKFSLR-----ENPVETKG-EAPS 392
DB 345 DAFQFC--QKKVFPQAYG-----LRAKNLMDINTSGLLVHPASQNSGENLGCRKSA 396

.QY 393 DGNIEVKLCBOTEE 406
DB 397 DNPQESLMEETGE 410

```

RESULT 2

```

NFF2 HUMAN
ID NFF2 HUMAN STANDARD; PRT; 522 AA.
AC Q9Y5X5; Q96RV1; Q9NR49;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Neuropeptide FF receptor 2 (Neuropeptide G protein-coupled receptor)
DE (G-protein-coupled receptor HLWAR77).
GN GPR74 OR NPGPR OR NPF2.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND TISSUE SPECIFICITY.
RC TISSUE=Fetal.
RX MEDLINE=99180505; PubMed=10079187;
RA Ciklos S., Gregor P., Koppal J.;
RT "Sequence and tissue distribution of a novel G-protein-coupled

```

```

receptor expressed prominently in human placenta.";
Biochem. Biophys. Res. Commun. 256:352-356(1999)).
[2]
SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
TISSUE=Brain;
MEDLINE=20408933; PubMed=10851242;
RA Eshourbagy N.A., Ames R.S., Fitzgerald L.R., Foley J.J., Buckley P.T.,
RA Chambers J.K., Szekeres P.G., Evans N.A., Schmidt D.B., Tan K.B.,
RA Dytko G.M., Murdock P.R., Milligan G., Groarke D.A., Wilson S.,
RA Shabou U., Nuthulaganti P., Wang D.Y., Bergsma D.J.,
RA Sarau H.M.;
RT "Receptor for the pain modulatory neuropeptides FF and AF is an orphan
RT G-protein-coupled receptor.";
J. Biol. Chem. 275:25965-25971(2000).
[3]
SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
TISSUE=Spinal cord;
MEDLINE=20564301; PubMed=11024015;
RA Bonini J.A., Jones K.A., Adham N., Forray C., Artymyshyn R.,
RA Durkin M.M., Smith K.E., Tamm J.A., Boteju L.W., Lakhani P.P.,
RA Raddatz R., Yao W.-J., Ogozalek K.B., Boyle N., Kouranova E.V.,
RA Quan Y., Vayesse P.J., Wetzel J.M., Branchek T.A., Gerald C.,
RA Borowsky B.;
RT "Identification and characterization of two G protein-coupled
RT receptors for neuropeptide FF.";
J. Biol. Chem. 275:39324-39331(2000).
[4]
SEQUENCE FROM N.A. (ISOFORM 3).
MEDLINE=2029143; PubMed=10837915;
RA Parker R.M.C., Copeland N.G., Eyre H.J., Liu M., Gilbert D.J.,
RA Crawford J., Couzens M., Sutherland G.R., Jenkins N.A., Herzog H.;
RT "Molecular cloning and characterization of GPR74 a novel G-protein
RT coupled receptor closest related to the Y-receptor family.";
Brain Res. Mol. Brain Res. 77:199-208(2000).
[5]
SEQUENCE FROM N.A. (ISOFORM 2).
RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA Wang R., Evans J., Gould R., Austin C.P.;
RT "Identification and characterization of two cognate receptors for
RT mammalian FMRFamide-like neuropeptides.";
Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
[6]
SEQUENCE FROM N.A. (ISOFORM 4).
RA Laemmle B.S., Schindler M., Beilmann M., Hamilton B.S., Doods H.N.,
RA Wieland H.A.;
RT "Cloning and characterization of the NPGR receptor and identification
RT of a novel short mRNA isoform in human hypothalamus.";
Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: RECEPTOR FOR NPFF (A-18-P-AMIDE) AND NPFF (P-8-P-AMIDE)
CC ALSO BE ACTIVATED BY A VARIETY OF NATURALLY OCCURRING PEPTIDES. CAN
CC FMRF-AMIDE LIKE LIGANDS. THIS RECEPTOR MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM.
CC SUBCELLULAR LOCATION: Integral membrane protein.
CC ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Comment=Experimental confirmation may be lacking for some
CC isoforms;
CC Name=1; Synonyms=long form;
CC IsoId=Q9Y5X5-1; Sequence=Displayed;
CC Name=2; Synonyms=short form;
CC IsoId=Q9Y5X5-2; Sequence=VSP_001907;
CC Name=3;
CC IsoId=Q9Y5X5-3; Sequence=VSP_001908, VSP_001909;
CC Name=4;
CC IsoId=Q9Y5X5-4; Sequence=VSP_001910, VSP_001911;
CC TISSUE SPECIFICITY: Isoform 1 is abundant in placenta. Relatively
CC highly expressed in thymus, testis, and small intestine. Expressed
CC at low levels in several tissues including spleen, prostate,
CC brain, heart, ovary, colon, kidney, lung, liver and pancreas and
CC not expressed in skeletal muscle and leukocytes. Highest but

```


relatively low level of isoform 2 in placenta and very low level in numerous tissues including adipose tissue and many brain regions. Isoform 3 is expressed in brain and heart and, at lower levels, in kidney, liver, lung and pancreas.

- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

- HIGHEST TO OREXIN, NPY AND CHOLECYSTOKININ RECEPTORS.

- CAUTION: Ref.4 sequence differs from that shown due to a frameshift in position 503.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

EMBL; AF119815; AAD2047.1; -
 DR EMBL; AF257210; AAG87078.1; -
 DR EMBL; AF268899; AAG41398.1; -
 DR EMBL; AF236083; AAK58513.1; ALT_FRAME.
 DR EMBL; AF330053; AAK94197.1; -
 DR EMBL; AJ311393; CAC85427.1; -
 DR Genew; HGNC:4525; GPR74.
 DR MIM; 607449; -
 DR GO; GO:0005897; C: integral to plasma membrane; TAS.
 DR GO; GO:0008151; P: cell growth and/or maintenance; TAS.
 DR GO; GO:0007186; P: G-protein coupled receptor protein signalin. .; TAS.
 DR GO; GO:0009582; P: perception of abiotic stimulus; TAS.
 DR InterPro; IPR0020276; GPCR_Rhodopsn.
 DR InterPro; IPR005395; NPFF_receptor.
 DR InterPro; IPR005397; NPFF_receptor2.
 DR Pfam; PF00001; 7tm.1; 1.
 DR PRINTS; PR00237; GPCRHOPOFSN.
 DR PRINTS; PR01570; NPFFRECEPTOR.
 DR PRINTS; PR01572; NPFFRECEPTOR.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Alternative splicing.
 FT DOMAIN 1 147 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 148 168 1 (POTENTIAL).
 FT DOMAIN 169 184 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 185 205 2 (POTENTIAL).
 FT DOMAIN 206 221 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 222 242 3 (POTENTIAL).
 FT DOMAIN 243 262 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 263 283 4 (POTENTIAL).
 FT DOMAIN 284 319 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 320 340 5 (POTENTIAL).
 FT DOMAIN 341 377 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 378 398 6 (POTENTIAL).
 FT DOMAIN 399 413 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 414 434 7 (POTENTIAL).
 FT DOMAIN 435 522 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 220 308 BY SIMILARITY.
 FT CARBOHYD 110 110 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 122 122 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 300 300 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 1 102 Missing (in isoform 2).
 FT FT /FTid=VSP_001907.
 FT VARSPLIC 1 99 Missing (in isoform 3).
 FT FT /FTid=VSP_001908.
 FT VARSPLIC 100 100 R -> M (in isoform 3).
 FT FT /FTid=VSP_001909.
 FT VARSPLIC 101 132 FINMEKWDGSSSEKHWFIWNNYKHLYSDI -> MAIWK
 FT FT HDVQDQWIGPNCISFSLYVSNCCR (in isoform
 FT FT 4).
 FT FT /FTid=VSP_001910.
 FT VARSPLIC 133 522 Missing (in isoform 4).
 FT FT /FTid=VSP_001911.
 FT VARSPLIC 466 466 A -> T (IN REF. 1 AND 4).

SQ SEQUENCE 522 AA; 60270 MW; 40CB9FCD42F77041 CRC64;
 Query Match 22.58; Score 502; DB 1; Length 522;
 Best Local Similarity 30.1%; Pred. No. 2.5e-28;
 Matches 128; Conservative 87; Mismatches 162; Indels 48; Gaps 12;

Qy 14 LLRDHNLTRQOFIALYELRLPLVYTPGLPGRALKALVLTGLIFALALFGNALVFYVVTSS 73
 Db 128 LYSIDNIT---VYNYLHQVQ-----AAFIISYFLIFLGMGNTVVCFIWNRN 175
 Qy 74 KAMRTVNIIFCSIALSLLITFCIPVTMLQNSDNWLGAFICKVVPFVQSTAVVTEI 133
 Db 176 KHMHTVNLFILALISDLLVIGIFCMPTITLDDNIAGWPGFNTMCKISGLVQGISVAASV 235
 Qy 134 LTWTCTAVERHQGLVNPFKMKQYTNERRAFTMLGVWLVAIVGSP---MHWVQOLBIKY 190
 Db 236 FTLVAVADVPQCVVPFVKPK--LTIKTAFVIIMIVLAIMTSPSAVMLHVOE-EKYY 292
 Qy 191 DFLYEKEH-----ICCLEEWTSPVHQKIYTTFTILVILFLPLVMVLLIYSKIGYELWIKK 245
 Db 293 RVLNSQNKTSPPVYWCREDWPNQEMRKIYTTVLPAHYLAPLSLIIVIMYGRIGISLFP--- 349
 Qy 246 RVGDGSLVRLTHGKMSK-----IARKKKRAVIMVMTVVALFAVCWAPFHVHVMMIEVSNF 301
 Db 350 ---RAAVPHT--GRKNQEQWHVVSRRKKQIKIKLLIALLFILSWLPLTLMMLSDVADL 404
 Qy 302 EKEYDDVTIKMIFAIQIIGFSNSICNPVIYAFMNFKNVLSAVCYCIVNKTFSQAOR 361
 Db 405 SPNELQIINIYIYPPFAHWLAFGNSSVNPILYGFNFENFRGFOEAFOLQCKRAKPMEA 464
 Qy 362 HGNSGITMMRKAKFSLRENFEVETKGEA-FSDGNIEVKLECEQTEEEKKLRHLALPRSE 420
 Db 465 YA-----LKAKSHVLINTSNQLVQESTFQNPFGHETLLYRKAEPQOE-----LVWEE 512
 Qy 421 LAENS 425
 Db 513 LKETT 517

RESULT 3
 NY2R_HUMAN STANDARD; PRT; 381 AA.
 ID NY2R_HUMAN Q13281; Q13457; Q9UEB7;
 AC P49146; Q13281; Q13457; Q9UEB7;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
 GN NPY2R.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Hippocampus;
 RX MEDLINE=96070760; PubMed=7592910;
 RA Gerald C., Walker M.W., Vaysse P.J.-J., He C., Branchek T.A.,
 RA Weinshank R.L.;
 RT "Expression cloning and pharmacological characterization of a human
 RT hippocampal neuropeptide Y/peptide YY Y2 receptor subtype.";
 RT J. Biol. Chem. 270:26758-26761(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96226058; PubMed=8632753;
 RA Gehlert D.R., Beavers L.S., Johnson D., Gackenhimer S.L.,
 RA Schober D.A., Gadsdi R.A.;
 RT "Expression cloning of a human brain neuropeptide Y Y2 receptor.";
 RT Mol. Pharmacol. 49:224-228(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96032678; PubMed=7559383;
 RA Rose P.M., Fernandes P., Lynch J.S., Frazier S.T., Fisher S.M.,

RA Kodukula K., Kienzie B., Seethala R.;
 RT "Cloning and functional expression of a cDNA encoding a human type 2
 RL neurotrophin Y receptor.";
 RN J. Biol. Chem. 270:22661-22664 (1995).
 [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96209788; PubMed=8643460;
 RA Yan H., Yang J., Marasco J., Yamaguchi K., Brenner S., Collins F.,
 RA Karbon W.;
 RT "Cloning and functional expression of cDNAs encoding human and rat
 RT pancreatic polypeptide receptors.";
 RL Proc. Natl. Acad. Sci. U.S.A. 93:4661-4665 (1996).
 [5]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=97131518; PubMed=8975716;
 RA Ammar D.A., Eadie D.M., Wong D.J., Ma Y.-Y., Kolakowski L.F. Jr.,
 RA Yang-Peng T.L., Thompson D.A.;
 RT "Characterization of the human type 2 neurotrophin Y receptor gene
 RT (NPY2R) and localization to the chromosome 4q region containing the
 RT type 1 neurotrophin Y receptor gene.";
 RL Genomics 38:392-398 (1996).
 [6]

RN SEQUENCE FROM N.A.
 RP Zastawny R.L.;
 RT "Human neurotrophin Y Y2 receptor gene.";
 RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE RANK
 CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS
 CC PYY > NPY > PYY (3-36) > NPY (2-36) > [ILE-31, GLN-34]. PP >
 CC [LEU-31, PRO-34] NPY > PP, [PRO-34] PYY AND NPY FREE ACID.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: HIGH LEVELS IN AMYGDALA, CORPUS CALLOSUM,
 CC HIPPOCAMPUS AND SUBTHALAMIC NUCLEUS. ALSO DETECTABLE IN CAUDATE
 CC NUCLEUS, HYPOTHALAMUS AND SUBSTANTIA NIGRA.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININ RECEPTORS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch.
 CC -----

DR EMBL; U36269; AAC50281.1; -
 DR EMBL; U42766; AAB04120.1; -
 DR EMBL; U32500; AAB03170.1; -
 DR EMBL; U42389; AAB07760.1; -
 DR EMBL; U50146; AAC51115.1; -
 DR EMBL; U76254; AAD00248.1; -
 DR PIR; I39187; I39187.
 DR Genes; HGNC:7957; NPY2R.
 DR MIM; 162642; -
 DR GO; GO:0005887; C: integral to plasma membrane; TAS.
 DR GO; GO:0005246; F: calcium channel regulator activity; TAS.
 DR GO; GO:0004983; F: neurotrophin Y receptor activity; TAS.
 DR GO; GO:0004872; F: receptor activity; TAS.
 DR GO; GO:0007193; P: G-protein signaling, adenylylate cyclase inhibi. . .; TAS.
 DR GO; GO:0007626; P: locomotory behavior; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm1.1; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1.1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1.2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate
 FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 50 72 1 (POTENTIAL).
 FT DOMAIN 73 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 104 2 (POTENTIAL).
 FT DOMAIN 105 124 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 125 146 3 (POTENTIAL).

FT DOMAIN 147 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 187 4 (POTENTIAL).
 FT DOMAIN 188 214 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 215 240 5 (POTENTIAL).
 FT DOMAIN 241 268 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 269 291 6 (POTENTIAL).
 FT DOMAIN 292 304 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 305 328 7 (POTENTIAL).
 FT DOMAIN 329 381 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 123 203 BY SIMILARITY.
 FT LIPID 342 342 PALMITATE (POTENTIAL).
 FT CONFLICT 134 134 Y -> A (IN REF. 3).
 FT CONFLICT 172 172 A -> G (IN REF. 4).
 FT CONFLICT 174 174 G -> R (IN REF. 4).
 FT CONFLICT 202 202 A -> P (IN REF. 4).
 FT CONFLICT 209 209 G -> A (IN REF. 4).
 FT CONFLICT 248 248 N -> S (IN REF. 6).
 FT CONFLICT 311 311 H -> Y (IN REF. 6).
 SQ SEQUENCE 381 AA; 42731 MW; 7D018C0169597BC7 CRC64;

Query Match 21.0%; Score 469.5; DB 1; Length 381;
 Best Local Similarity 31.0%; Pred. No. 3.5e-26;
 Matches 113; Conservative 75; Mismatches 123; Indels 53; Gaps 13;
 QY 38 PELPGRKML-----ALVLTGVLFALALFGNALVYVYVTRSKAMRTVTNIFICSLALS 92
 DB 38 PELDSTKLEIVQVVLVLAICSIILLGVNLSLVHVIKFSMTVTNFIANLAVADL 97
 QY 93 LITFCIPVTMLQNIISNMLGGAFICRQVFPVQSTAVVTILMTCTIAVERHQGLVHPFK 152
 DB 96 LVNTLCLEPFTLTTLTLMGEMKGPVLCPLVPAQGLAVQVSTITLTVALDRHRCIVVHLE 157
 QY 153 MKQVYTNRAFTMLGVVNLVAVIVGSPMWHVQV---LEIKYDLYEKEHICCLPEW---T 206
 DB 158 SK--ISKREISFLIIGLAWGISALLASPLAFREVSLEIIPDF-----EIVACTEKWPEE 211
 QY 207 SPVHQKIYTTTILFLLPLMVMMLIYKIGYELW--IKRVRGDSGLVRLTHGKMSKI 264
 DB 212 KSIYGVTVYSLSLILYVPLGIISFSYTRI---WSKLNHVSPGAANDHYH----- 260
 QY 265 ARKKRAVIMVMTVVALPANCWAFHVVHVMIEYSNFE---KEYDDVTIKMIPAIVOLIG 321
 DB 261 -ORRQTKMLVCVVVFAVSWLPLHAPQLAVDIDSQVLDLKEY-----KLIFTVFHIIA 314
 QY 322 PENSICNPVYAFNMENFKNVLGAVCYCIVNKFSPAQK----HGNSGITMERKAKFS 377
 DB 315 MCSTFANPLLYGMNNSYRKAFLSA-----FRCEQLDAIHSEVSVTFKAK-NLE 364
 QY 378 LREN 381
 DB 365 VRKN 368

RESULT 4
 NY2R_MACMU STANDARD; PRT; 381 AA.
 AC Q9GK74;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
 GN NPY2R.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopitheinae; Macaca.
 OX NCBI_TaxID=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21184974; PubMed=11287088;
 RA Gehlert D.R., Yang P., George C., Wang Y., Schober D.,
 RA Gackenhimer S., Johnson D., Beavers L.S., Gadske R.A., Baez M.;


```

RT RT Cloning and characterization of Rhesus monkey neuropeptide Y receptor
RL Peptides 22:343-350 (2001).
CC -|- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF303090; AAG40772.1; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHOPOPSN.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 50 72 1 (POTENTIAL).
CC DOMAIN 73 82 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 83 104 2 (POTENTIAL).
CC DOMAIN 105 124 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 125 146 3 (POTENTIAL).
CC DOMAIN 147 166 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 167 187 4 (POTENTIAL).
CC DOMAIN 188 214 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 215 240 5 (POTENTIAL).
CC DOMAIN 241 268 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 269 291 6 (POTENTIAL).
CC DOMAIN 292 304 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 305 328 7 (POTENTIAL).
CC DOMAIN 329 381 CYTOPLASMIC (POTENTIAL).
CC CARBOHYD 11 11 N-LINKED (GLCNAC. .) (POTENTIAL).
CC DISULFID 123 203 BY SIMILARITY.
CC LIPID 342 342 PALMITATE (POTENTIAL).
CC SEQUENCE 381 AA; 42729 MW; D689D691DC05A49D CRC64;
Query Match
Best Local Similarity 30.8%; Score 468.5; DB 1; Length 381;
Matches 112; Conservative 76; Mismatches 123; Indels 53; Gaps 13;
QY 38 PELPGRAKL-----ALVLTGLVLIIPALFGNALFYVYVTRSKAMRTVNTNIFCSLALSDL 92
Db 38 PELIDTKLIEVQVVLILAYCSIIILGVIGSLVHVIVKPSKMTVTNPFIALAVADL 97
QY 93 LITFFCIPVTMLQNISDNWLGAPICRMVPFVQSTAVVTEILTMTCLAVRHQGLVHPFK 152
Db 98 VVNTLCPLFTLTLYLMGEKMGKPVLCVLPVPAQGLAVQVSTITLTVALDHRICVHLE 157
QY 153 MKQWYTRRAFTMLGVVWLVAIVGSPMHWVQ---LEIKYDFLYEKEHICCLSEW---T 206
Db 158 SK--ISKRIISFLITGLAWGISALLASPLAIFREYSLIEIIPDF---EIVACTEKWPCEE 211
QY 207 SPVHQKTYTTFILVILFLPLMWMLLYSKIGYELW--IKKRVGDGSGVLRITGKEMSKI 264
Db 212 KSIYGVTVISLSSLLILVPLGIISFSYTRT---WSKLKSHVSPGANDHYH----- 260
QY 265 ARKKRAVIMVTVVVALFAVCWAPFHVHVMMIYSNFE---KEYDDVTIKMIFAIVQIG 321
Db 261 -QRQKTKKLVCVVVVFAVSNLPLHAFQLAVDIDSHVLDKEY-----KLITVFVHIA 314
QY 322 PNSSICNPVIVAFNENFKKNVLGAVCYIVNKTFTSPAQR-----HGNSGITMWRKAKFS 377
Db 315 MCSTFANPLLYGWNNSVYKAFLSA-----FRCEQRDLATHSEVSTFRKAK-NLE 364
QY 378 LREN 381

```

```

Db 365 VRKN 368
:|:|
NY2R PIG STANDARD; PRT; 382 AA.
RESULT 5
ID NY2R_PIG STANDARD; PRT; 382 AA.
AC O02836; Q9TS11;
DT 15-JUL-1998 (Rel. 36, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=99017380; PubMed=9802394;
RA Malmstrom R.E., Hoekfelt T., Bjoerkman J.-A., Nihlen C., Bystroem M.,
RA Ekstrand A.J., Lundberg J.M.;
RT "Characterization and molecular cloning of vascular neuropeptide Y
RT receptor subtypes in pig and dog.";
RL Regul. Pept. 75:55-70(1998).
RN [2]
RP REVISION TO 207.
RA Ekstrand A.J.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wraith A., Tornsten A., Chardon P., Harbitz I., Chowdhary B.P.,
RA Andersson L., Larhammar D.;
RT "Porcine NPY receptors NPY1R, NPY2R and NPY5R: cloning, mapping and
RT comparative analysis.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -|- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -|- SUBCELLULAR LOCATION: Integral membrane protein.
CC -|- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF005780; AAC26670.2; -
CC EMBL; AF106082; AAD13777.1; ALT_INIT.
CC HSPP; P02699; 1F88.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHOPOPSN.
CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL 1; 1.
CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL 2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Phosphorylation; Lipoprotein; Palmitate.
CC DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 51 73 1 (POTENTIAL).
CC DOMAIN 74 83 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 84 105 2 (POTENTIAL).
CC DOMAIN 106 125 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 126 147 3 (POTENTIAL).
CC DOMAIN 148 167 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 168 188 4 (POTENTIAL).
CC DOMAIN 189 215 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 216 241 5 (POTENTIAL).
CC DOMAIN 242 269 CYTOPLASMIC (POTENTIAL).
CC TRANSMEM 270 292 6 (POTENTIAL).
CC DOMAIN 293 EXTRACELLULAR (POTENTIAL).

```


RT "To eat or to sleep? Orexin in the regulation of feeding and
 RL wakefulness.";
 CC -!- FUNCTION: NONSELECTIVE, HIGH-AFFINITY RECEPTOR FOR BOTH OREXIN-A
 CC AND OREXIN-B NEUROPEPTIDES.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AK038551; BAC30039.1; -;
 CC DR EMBL; AK048781; BAC33457.1; -;
 CC DR EMBL; AF394597; AAK71327.1; -;
 CC DR MGD; MGI:1889024; Mox2r.
 CC DR InterPro; IPR000276; GPCR_Rhodopsn.
 CC DR InterPro; IPR004060; Orexin_receptor2.
 CC DR Pfam; PF00001; 7tm 1; 1
 CC DR Pfam; PF03827; Orexin_rec2; 1.
 CC DR PRINTS; PR00237; GPCR_Rhodopsn.
 CC DR PRINTS; PR01522; OREXIN2R.
 CC DR PROSITE; PS00237; G-PROTEIN RECP FL 1; 1.
 CC DR PROSITE; PS0262; G-PROTEIN RECP FL 2; 1.
 CC DR G-protein coupled receptor; Transmembrane; Glycoprotein.
 KW DOMAIN 1 54
 FT DOMAIN 1 54
 FT TRANSMEM 55 75
 FT DOMAIN 76 88
 FT DOMAIN 89 110
 FT TRANSMEM 111 127
 FT DOMAIN 128 150
 FT TRANSMEM 151 172
 FT DOMAIN 173 193
 FT TRANSMEM 194 224
 FT DOMAIN 225 247
 FT TRANSMEM 248 304
 FT DOMAIN 305 327
 FT TRANSMEM 328 342
 FT DOMAIN 343 366
 FT TRANSMEM 367 460
 FT CARBOHYD 14 14
 FT CARBOHYD 22 22
 FT CARBOHYD 202 202
 FT CONFLICT 201 201
 FT CONFLICT 240 240
 FT CONFLICT 240 240
 FT SEQUENCE 460 AA; 52461 MW; D62A67C15BA67DCC CRC64;
 SQ
 Query Match 20.6%; Score 460.5; DB 1; Length 460;
 Best Local Similarity 26.0%; Pred. No. 1.8e-25;
 Matches 118; Conservative 97; Mismatches 164; Indels 75; Gaps 11;
 QY 4 LNITEQFSRLDRNLNTEQFIAYLRPLVYTFELPGRAKALVLGLVLPALALPGN 63
 Db 21 LNTEQFP---LNPTDYDDEEFLR-YLMREYLL-PRYEWVLIAGVILVFFVVALIGN 72
 QY 64 ALVFVVTTRSKAMRVTVNIFICSLALSDLLITFFCIPVTVMLNIDNMLGAFICKVYVF 123
 Db 73 VLVCVAVKNNHMRVTVNIFVNLADVLVITCLPDLVVDITETFPFGOSLCKVIPY 132
 QY 124 VQSTAVVTEILTMTCIAVERHGLVHPFKMKQWYTNRRAFVTLGVVNLVAVIVGSPMWHV 183
 Db 133 LQTVSVSVSVLTSLCALDRWVAICHLPMFK--STAKRARSIVIVIVSVCLIMIPQAV 190
 QY 184 QOLEIKYDFLYEKEH--CCLEWTSPTVKQIYTFILVFLPLMNLILYSKIGVEL 241
 Db 191 MECSSMLPGLANKTTLFTVCDHGWGVEYPKMHCFFELVTYMAPLCIMILAYIQLPRKL 250
 QY 242 WIKKRVGDGSLR-----THGKMSKIARKKRAVIMMVTV 278

Db 251 WCRQIPGTSSVQKWKQKQPVQSRGQQSKARISAVAAEIKQIRARKTARMLMV-V 309
 QY 279 VALFVAVWAPFHVVMHIE-YSNFKEYDDVTIKMFAIVQIGFNSNCINPIVYAFMNE 337
 Db 310 LLVFAICLPIISILNVLKRVFGMFTHTEDRETVAWFTPSHVLVYANGSAANPIYNFLSG 369
 QY 338 NFKKNVLGAVCVIVNKTFPSPAQRHNGSITMWRKKAKFSLRENVEETKGAFGSDGNIE 397
 Db 370 KREFEFKAFAFSCCL-----GVHRRQGDRLARGR-- 397
 QY 398 VKLCBQTEKKKLRHLALFR--SELAENSPIDS 429
 Db 398 ----TSTSRKSLTTQISNFDNVSKLSEHVILTS 427
 RESULT 9
 ID OX2R RAT STANDARD; PRT; 460 AA.
 AC P56719;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
 GN HCRTR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98150861; PubMed=9491897;
 RA Sakurai T., Amemiya A., Ishii M., Matsuzaki I., Chemelli R.M.,
 Tanaka H., Williams S.C., Richardson J.A., Kozlowski G.P., Wilson S.,
 Arch J.R.S., Buckingham R.E., Haynes A.C., Carr S.A., Annan R.S.,
 McNulty D.E., Liu W.-S., Terrett J.A., Elshourbagy N.A., Bergsma D.J.,
 Yanagisawa M.;
 RT "Orexins and orexin receptors: a family of hypothalamic neuropeptides
 and G protein-coupled receptors that regulate feeding behavior.";
 RL Cell 92:573-585 (1998).
 RN [2]
 RP REVIEW.
 RX MEDLINE=21237974; PubMed=11340621;
 RA Hungs M., Mignot E.;
 RT "Hypocretin/orexin, sleep and narcolepsy.";
 RL Bioessays 23:397-408 (2001).
 RN [3]
 RP REVIEW.
 RX MEDLINE=21178476; PubMed=11283317;
 RA Willie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
 RT "To eat or to sleep? Orexin in the regulation of feeding and
 wakefulness.";
 RL Annu. Rev. Neurosci. 24:429-458 (2001).
 CC -!- FUNCTION: NONSELECTIVE, HIGH-AFFINITY RECEPTOR FOR BOTH OREXIN-A
 CC AND OREXIN-B NEUROPEPTIDES.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed in the brain in the cerebral
 CC cortex, septal nuclei, hippocampus, medial thalamic nuclei including
 CC and median raphe nuclei, and many hypothalamic nuclei including
 CC the tuberomammillary nucleus, dorsomedial hypothalamus,
 CC paraventricular hypothalamic nucleus, and ventral premammillary
 CC nucleus. Not detected in the spleen, lung, liver, skeletal muscle,
 CC kidney and testis. Orexin receptor mRNA expression has also been
 CC reported in the adrenal gland, enteric nervous system, and
 CC pancreas.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

```

CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL; AF041246; AAC04042.1; -.
CC      InterPro; IPR000276; GPCR_Rhodpsn.
CC      InterPro; IPR004060; Orexin_receptor2.
CC      Pfam; PF00001; 7tm_1; 1.
CC      Pfam; PF03827; Orexin_rec2; 1.
CC      PRINTS; PRO0237; GPCRHHODPSN.
CC      PRINTS; PRO1522; OREXIN2R.
CC      PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
CC      PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.
CC      G-protein coupled receptor; Transmembrane; Glycoprotein.
CC      DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).
CC      TRANSSEM 55 75
CC      DOMAIN 76 88
CC      TRANSSEM 89 110
CC      DOMAIN 111 127
CC      TRANSSEM 128 150
CC      DOMAIN 151 172
CC      TRANSSEM 173 193
CC      DOMAIN 194 224
CC      TRANSSEM 225 247
CC      DOMAIN 248 304
CC      TRANSSEM 305 327
CC      DOMAIN 328 342
CC      TRANSSEM 343 366
CC      DOMAIN 367 460
CC      CARBOHYD 14 14
CC      TRANSSEM 22 22
CC      CARBOHYD 202 202
CC      SEQUENCE 460 AA; 52489 MW; 3844E3D8F88B5D5 CRC64;

Query Match
Best Local Similarity 20.6%; Score 459.5; DB 1; Length 460;
Matches 118; Conservative 97; Mismatches 164; Indels 75; Gaps 11;

QY      4 LNITPEPSRLRLDHLNLTREOFIALYRLPLVYTFELPGRAKLALVTGLVLFALAFGN 63
DB      21 LNITQEPF---LNPTDXDDEEFLR-YLWREYLH----PKYEWVLVAGYIIVFVVALIGN 72
QY      64 ALVYVYVTRSKAMTVTNIFCSLSDLLATFFCIPVTMLQNSDNLWGLGAFCKMVPF 123
DB      73 VLVCVAVKNNHMTVNTVNFVNLSDLVLTITCLPATLVVDITETWFFQOSLCKVIPY 132
QY      124 VQSTAVTEILTMTCTIAVERHQGLVHPFKMKWQVYNNRAFTMLGWLVLVAVIGSPMHV 183
DB      133 LQTVSVSVSLTSCIALDRVVAICHPLMFK--STAKPARNISVIVWIVSCIIIPQAV 190
QY      184 QOLEIKYDFLYEKHI--CCLEWTSVHQKIYTFILVILFLPLMVMMLYLSKIGYEL 241
DB      191 MERSMLPLGLANKTTLFTVCDERWGEVYPKVYHICFELVYMAPLCLMWLAYLQIFKL 250
QY      242 WIKKRVGDSGLR-----TIHGKEMSKIARKKRAVIMVTV 278
DB      251 WCRIPGTSSVQKRWKQVQPSQPRGSGQSKARISAAAEIKQIRARRKTAMLMV-V 309
QY      279 VALFVACWAPVHVHMMIE--YSNFEKEDVDVTIKMFAIVQIIGFSNCSNPIVAFWNE 337
DB      310 LLVFAICVPLPISILNLKRVFGMTHTEDRETVAWFTFHSWLVAVYNSAANPIIYNLSG 369
QY      338 NFKKNVLISAVCYIVNKTFTSPAQRHSGSGITMRRKAKFSLRENPVETKGEAFSDGNI 397
DB      370 KFEFKAARFSCCL-----GVHRRQDRLARGR-- 397
QY      398 VKLCEQTEKKKLKRLHALPR--SELAENSPLDS 429
DB      398 ----TSTESKSLTQISFNDFNVSKLSEHALTS 427

RESULT 10
ID_NY2R_CAVPO STANDARD; PRT; 381 AA.
AC Q922D5;

```

```

DT      30-MAY-2000 (Rel. 39, Created)
DT      30-MAY-2000 (Rel. 39, Last sequence update)
DE      30-MAY-2000 (Rel. 39, Last annotation update)
DE      Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN      NPY2R.
OS      Cavia porcellus (Guinea pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Euthera; Rodentia; Hystricognathi; Caviidae; Cavia.
OX      NCB TaxID=10141;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=99017376; PubMed=9802390;
RA      Sharma P.S., Holmberg S.K., Eriksson H., Beck-Sicking A.G.,
RA      Grundemar L., Larhammar D.;
RA      "Cloning and functional expression of the guinea pig neuropeptide Y
RT      Y2 receptor.";
RL      Regul. Pept. 75:23-28(1998).
CC      -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC      -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC      -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC      HIGHEST TO TACHYKININS RECEPTORS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
CC      EMBL; AF072821; AAD13143.1; -.
CC      InterPro; IPR000276; GPCR_Rhodpsn.
CC      Pfam; PF00001; 7tm_1; 1.
CC      PRINTS; PRO0237; GPCRHHODPSN.
CC      PROSITE; PS00237; G-PROTEIN RECP_F1_1; 1.
CC      PROSITE; PS00262; G-PROTEIN RECP_F1_2; 1.
CC      G-protein coupled receptor; Transmembrane; Glycoprotein;
CC      Phosphorylation; Lipoprotein; Palmitate.
CC      DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
CC      TRANSSEM 51 73
CC      DOMAIN 74 83
CC      TRANSSEM 84 105
CC      DOMAIN 106 125
CC      TRANSSEM 126 147
CC      DOMAIN 148 167
CC      TRANSSEM 168 188
CC      DOMAIN 189 215
CC      TRANSSEM 216 241
CC      DOMAIN 242 269
CC      TRANSSEM 270 292
CC      DOMAIN 293 305
CC      TRANSSEM 306 329
CC      DOMAIN 330 381
CC      CARBOHYD 11 11
CC      TRANSSEM 124 204
CC      TRANSSEM 343 343
CC      SEQUENCE 381 AA; 42342 MW; D806859A43EAC CRC64;

Query Match
Best Local Similarity 20.4%; Score 457; DB 1; Length 381;
Matches 108; Conservative 75; Mismatches 123; Indels 50; Gaps 12;

QY      38 PELPGRAKL-----ALVLTGLVLFALAFGNALVFVWTFKAMRVTNIFCSLSD 92
DB      39 PELIDSTKLTETVRVLIILAYCSIIILGVGNSLVHVIKFKSMRTVTFNFIANLAVAD 98
QY      93 LIITFCIPVTMLQNSDNLWGLGAFICKMVPFVQSTAVVTEILTMTCTIAVERHQGLVHPFK 152
DB      99 LVNLTCLPFTLYTLMGKWKMGVFLCHLPYAGLAVQVSTVTLTVALDRHRCIVHLD 158
QY      153 MKVQYNNRAFTMLGVVWLVAVIGSPMHVQO---LEIKYDFLYEKHIKCLEW---T 206
DB      159 SK--ISKQNSFLIIGLAWGISALLASFLAIFREYSLIEIIPDF---EIVACTEKWPGEE 212

```


DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Orexin receptor type 2 (Ox2r) (Hypocretin receptor type 2).
 GN HCRTR2.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN (1)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93385793; PubMed=10458611;
 RA Lin L., Faraco J., Li R., Kadotani H., Rogers W., Lin X., Qiu X.,
 RA de Jong P.J., Nishino S., Mignot E.;
 RT "The sleep disorder canine narcolepsy is caused by a mutation in the
 RT hypocretin receptor 2 gene.";
 RL Cell 98:365-376(1999).
 RN (2)
 RP REVIEW.
 RX MEDLINE=21237974; PubMed=11340621;
 RA Hungs M., Mignot E.;
 RT "Hypocretin/orexin, sleep and narcolepsy.";
 RL Bioessays 23:397-408(2001).
 RN (3)
 RP REVIEW.
 RX MEDLINE=21178476; PubMed=11283317;
 RA Wallie J.T., Chemelli R.M., Sinton C.M., Yanagisawa M.;
 RT "To eat or to sleep? Orexin in the regulation of feeding and
 RT wakefulness.";
 RL Annu. Rev. Neurosci. 24:429-458(2001).
 RN (4)
 RP VARIANT NARCOLEPSY LYS-54, AND MUTAGENESIS OF GLU-54.
 RX MEDLINE=21180003; PubMed=11282968;
 RA Hungs M., Fan J., Lin L., Lin X., Maki R.A., Mignot E.;
 RT "Identification and functional analysis of mutations in the hypocretin
 RT (orexin) genes of narcoleptic canines.";
 RL Genome Res. 11:531-539(2001).
 CC -1- FUNCTION: NONSELECTIVE, HIGH-AFFINITY RECEPTOR FOR BOTH OREXIN-A
 CC AND OREXIN-B NEUROPEPTIDES.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DISEASE: Defects in HCRTR2 are a cause of an autosomal recessive
 CC form of narcolepsy, observed in Labrador, doberman and
 CC dachshunds. Narcolepsy is a neurological sleep disorder affecting
 CC animals and humans, characterized by excessive daytime sleepiness,
 CC sleep fragmentation, symptoms of abnormal rapid-eye-movement
 CC (REM) sleep, such as cataplexy, hypnagogic hallucinations, and
 CC sleep paralysis. Cataplexy is a sudden loss of muscle tone
 CC triggered by emotions which is the most valuable clinical feature
 CC used to diagnose narcolepsy. As in humans, most cases of canine
 CC narcolepsy are sporadic but an autosomal recessive form was also
 CC observed.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF164626; AAP49333.1; -
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR InterPro; IPR004060; Orexin_receptor2.
 DR Pfam; PF00001; 7tm.1; 1.
 DR Pfam; PF03827; Orexin_rec2; 1.
 DR PRINTS; PR00237; GPCR_Rhodopsn.
 DR PRINTS; PR01522; OREXIN2R.
 DR PROSITE; PS00237; G_PROTEIN_RECF1_1; 1.
 DR PROSITE; PS0262; G_PROTEIN_RECF1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Disease mutation.
 RN DOMAIN 1 54 EXTRACELLULAR (POTENTIAL).
 FT

FT TRANSMEM 55 75 1 (POTENTIAL).
 FT DOMAIN 76 88 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 89 110 2 (POTENTIAL).
 FT DOMAIN 111 127 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 128 150 3 (POTENTIAL).
 FT DOMAIN 151 172 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 173 193 4 (POTENTIAL).
 FT DOMAIN 194 224 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 225 247 5 (POTENTIAL).
 FT DOMAIN 248 307 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 308 327 6 (POTENTIAL).
 FT DOMAIN 328 342 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 343 366 7 (POTENTIAL).
 FT DOMAIN 367 444 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 14 14 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 202 202 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 54 54 E -> K (IN AUTOSOMAL RECESSIVE
 FT NARCOLEPSY).
 FT MUTAGEN 54 54 E -> K: LOSS OF FUNCTION.
 SQ SEQUENCE 444 AA; 50675 MW; D848A4536D485D6B CRC64;
 Query Match 20.2%; Score 451; DB 1; Length 444;
 Best Local Similarity 27.7%; Pred. No. 8.2e-25;
 Matches 104; Conservative 90; Mismatches 143; Indels 38; Gaps 9;
 QY 4 LNITPQFSLLRDHNLTRQFTALYELPLVVTPELPGRAKLALVLTGLVLPALFQ 63
 DB 21 LNITQEPF---LNPTDYDDDEFLR-VLWREYLH-----PKYEWVLIAGYIIVFVVALVGN 72
 QY 64 ALVFYVVTSSKAMRTVTNIFCSLALSDDLITFPFCIPVTMLQNISDNWLGAFICKMVPF 123
 DB 73 VLQCVAVKNNHMTVTNIVFVNLADLVLTITCLPATLVVDITETWFGQSICKVIPY 132
 QY 124 VQSTAVTEILTWCTIAVEHQGLVHPFKMKQYTNRRAFMTLGVVWLVAIVGVSPMHHV 183
 DB 133 LQTVSVSVSLTSLCDALDWAICHPLEK--STAKRANSIVIIWVSCIIMPOAIV 190
 QY 184 QQLLEIKYDELYEKEHI--CCLEEWTSVHOKIYTTFTLVILFLLPLMMLLYSKIGYEL 241
 DB 191 MEQSTMLPLGKANTLTFTVCDERWGEIYPMVHICFFLVTYMAPLCMLVLAQLQIFRKL 250
 QY 242 WIKRRVGDGSLVR-----TIHGKEMSKIARKKKRAVIMMVTY 278
 DB 251 WCQIPGTSVQVKWKPLQFASQPRGQGTQSRISAVAAEIKQIRARKRTARMLMV-V 309
 QY 279 VALPVCWAPFHHVHMIE-YSNFEKEYDDVTIKMIPAIIVQIIGFSNISCNPVIYAFMNE 337
 DB 310 LLVFAICYLPISILNVLKRVFGMFTHTEDRETVYAWFTFSGHVLVYANSAANPIYNFLSG 369
 QY 338 NPKKNVLSAV-CYCI 351
 DB 370 KFEFKAAPFSCCL 384
 RESULT 13
 NY4R_MOUSE
 ID NY4R_MOUSE STANDARD; PRT; 375 AA.
 AC Q61041;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DE Neuropeptide Y receptor type 4 (NPY4-R) (Pancreatic polypeptide
 DE receptor 1) (Pr1) (NPYR-D).
 GN PPYR1 OR NPY4R.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=129/SV;
 RC MEDLINE=96193913; PubMed=8641440;
 RX

RA Gregor P., Millham M.L., Feng Y., Decarr L.B., McCaleb M.L.,
 RA Cornfield L.J.;
 RT "Cloning and characterization of a novel receptor to pancreatic
 RT polypeptide, a member of the neuropeptide Y receptor family";
 RL FEBS Lett. 381:58-62(1996).
 CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE RANK
 CC ORDER OF AFFINITY OF THIS RECEPTOR FOR PANCREATIC POLYPEPTIDES IS
 CC PP >> PYY >= NPY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: HEART, DETECTED IN SMALL INTESTINE.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U40189; AAC52442.1; --
 CC PIR; S63685; S63685.
 CC MGI; MGI:105374; Ppy1.
 CC GO; GO:0001602; F:pancreatic polypeptide receptor activity; IDA.
 CC GO; GO:0001601; F:peptide YY receptor activity; IDA.
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
 CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 63 1 (POTENTIAL).
 FT DOMAIN 64 74 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 75 95 2 (POTENTIAL).
 FT DOMAIN 96 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 157 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 158 178 4 (POTENTIAL).
 FT DOMAIN 179 213 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 214 235 5 (POTENTIAL).
 FT DOMAIN 236 255 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 256 288 6 (POTENTIAL).
 FT DOMAIN 289 301 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 302 325 7 (POTENTIAL).
 FT DOMAIN 326 375 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 2 2 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 19 19 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 187 187 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 114 201 BY SIMILARITY.
 FT LIPID 340 340 PALMITATE (POTENTIAL).
 SQ SEQUENCE 375 AA; 42634 MW; DFFD31775DBA310E CRC64;
 Query Match 20.0%; Score 446.5; DB 1; Length 375;
 Best Local Similarity 30.6%; Pred. No. 1.4e-24;
 Matches 97; Conservative 72; Mismatches 121; Indels 27; Gaps 8;
 QY 46 LALVLTGVLI-FALALFGNALVFVYVTRSKMRVTNIFICSIALSDLLITFFCIPTVTL 104
 DB 41 LAFITTVSIEILGVGLNCLIFVTTRQKSNVNTLLIANLAFSDFMLICQPLVIT 100
 QY 105 QNISDNLGGAFICKVPFVQSTAVVTEILTTCIAVERHQGLVHPFKMKQYTNKRAFT 164
 DB 101 YTIMDYWIFGEVLCVMLTITQCMSTVTSLSLVLALEHQLIINP--TGWKPSIFQAVL 158
 QY 165 MGVVNLVAVIVGSP-----MHWVQOLEIKYDFLYEKEHI CCLBWTSPVHQKIYT 215
 DB 159 GIWVWIFSCFSLPFLANSTLNDLFFVNHRSKV-VVEFLEDK--VVCVFSWSSDHRRLIYT 215
 QY 216 TFIILVFLPLMVMVILYKIGYELWIKKRVGDSGLVTLTGKEMSKIARKKRAVIMM 275

Db 216 TFLLLFOYCIPLAFILVCYIRIQLOFQKHV-----FHAAACSSSAGOMKINSML 267
 QY 276 VTVVALFVWCWAPHVHVHMEIYSNFEXEYDDVIKMTFAIVQVIGFNSNCNPVIVAPM 335
 Db 268 MTWVTAFAVLPLPLHLVFNLTEDW--YQSAIPACHGNLFLMCHLLAMASTCVNPFYIGFL 325
 QY 336 NENFKKNVLSAV--CYC 350
 Db 326 NINFKDKIKALVLTCHC 342
 RESULT 14
 ID NY6R RABIT STANDARD; PRT; 371 AA.
 AC P79217;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Neuropeptide Y receptor type 6 (NPY6-R).
 GN NPY6R.
 OS Eukaryotagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Japanese white; TISSUE=Skeletal muscle;
 RX MEDLINE=97068888; PubMed=8910290;
 RA Matsumoto M., Nomura T., Momose K., Ikeda Y., Kondou Y., Akiho H.,
 RA Togami J., Kimura Y., Okada M., Yamaguchi T.;
 RT "Inactivation of a novel neuropeptide Y/peptide YY receptor gene in
 RT primate species";
 RL J. Biol. Chem. 271:27217-27226(1996).
 CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY. THE ACTIVITY
 CC OF THIS RECEPTOR IS MEDIATED BY G PROTEINS THAT INHIBITS ADENYLATE
 CC CYCLASE ACTIVITY.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC HIGHEST TO TACHYKININS RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D86521; BAA13104.1; --
 CC InterPro; IPR000276; GPCR_Rhodpsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GPCRHHODPSN.
 CC PROSITE; PS00237; G PROTEIN RECEPTOR FL1; 1.
 CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 34 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 35 57 1 (POTENTIAL).
 FT DOMAIN 58 69 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 70 90 2 (POTENTIAL).
 FT DOMAIN 91 110 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 111 132 3 (POTENTIAL).
 FT DOMAIN 133 152 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 153 174 4 (POTENTIAL).
 FT DOMAIN 175 208 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 209 230 5 (POTENTIAL).
 FT DOMAIN 231 261 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 262 284 6 (POTENTIAL).
 FT DOMAIN 285 297 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 298 321 7 (POTENTIAL).
 FT DOMAIN 322 371 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 11 N-LINKED (GLCNAC. . .) (POTENTIAL).


```
FT CARBOHYD 18 18 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 109 196 BY SIMILARITY.
FT LIPID 336 336 PALMITATE (POTENTIAL).
SQ SEQUENCE 371 AA: 42447 MW: 5847E358E320F24 CRC64;

Query Match
Best Local Similarity 20.0%; Score 446; DB 1; Length 371;
Matches 100; Conservative 93; Mismatches 117; Indels 52; Gaps 13;

QY 41 PGRALALVGLVLFALALFNG-ALVFFVTVTRSVAMRTVTVNIFCSLALSDLLITFCFI 99
Db 31 PSLALLLLLIAYTVVLLINGICNLSLTIIPKKQREARQNTVILLANLSLDILVCVMCI 90
QY 100 PVTMLQNTSDNWLGFATCKMVPFVQSTAVTEILTTCIAVERHOGVLVHPFKMKNQVTN 159
Db 91 PFTAIYTLMDRWIFGNTWCKLTYSVQSISVSIFSLVLIATIERQLVNP--RGWKPSA 148
QY 160 RRAFTMLGVWLVAVIVGSPM---WHV-----QQLKIKYFLYEKEHICCLLEMTSPVHQ 211
Db 149 SHAYWGIMLIWFLSLLSIPLLSYHLTDEPRNLSLPTD-LY-SHHVVVCVEHWPFSKTQ 206
QY 212 KIYTFILVILFLPLVWMLILYKIGYELWIKKVGSGVLRTHGKMSKIARKK--- 268
Db 207 LLYSTSLIMQYFVPLGFMFICYLKI-----VICLH-KRNSKIDRRRENE 250
QY 269 -----KRAVIMVMTVWALFAVCWAPFFVHMIEYSNFEKEDDVTIK---MFAIVQ 318
Db 251 SRLTENKINTMLISIVTFAACWLPNTFNVPD-----YHEVMSCHDLVFAICH 304
QY 319 IIGFNSNCNPIVAFNMENFKKNVLSAVCVIVNKTFTSPAQRHONGSI-TMRKXAKFS 377
Db 305 LVAMVSTCINPLFYGLFNLPKOLWLIHHCL---CFALRERYENIAISTLHTDESKGS 361
QY 378 LR 379
Db 362 LR 363

RESULT 15
NY2R_CHICK
ID NY2R_CHICK STANDARD; PRT; 385 AA.
AC Q9DDN6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neuropeptide Y receptor type 2 (NPY2-R) (NPY-Y2 receptor).
GN NPY2R.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20532533; PubMed=11078884;
RA Salanek E., Holmberg S.K., Berglund M.M., Boswell T., Larhammar D.;
RT "Chicken neuropeptide Y receptor Y2: structural and pharmacological
RT differences to mammalian Y2";
RL FEBS Lett. 484:229-234 (2000).
CC -!- FUNCTION: RECEPTOR FOR NEUROPEPTIDE Y AND PEPTIDE YY.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC HIGHEST TO TACHYKININS RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
```

```
DR EMBL; AF309091; AAG37898.1; -.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PRO0237; GPCRHHODOPSN.
DR PROSITE: PS00237; G-PROTEIN RECEPTOR FL 1; 1.
DR PROSITE: PS0262; G-PROTEIN RECEPTOR FL 2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Phosphorylation; Lipoprotein; Palmitate.
FT DOMAIN 1 53 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 54 76 1 (POTENTIAL).
FT DOMAIN 77 86 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 87 108 2 (POTENTIAL).
FT DOMAIN 109 128 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 129 150 3 (POTENTIAL).
FT DOMAIN 151 170 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 171 191 4 (POTENTIAL).
FT DOMAIN 192 218 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 219 244 5 (POTENTIAL).
FT DOMAIN 245 272 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 273 295 6 (POTENTIAL).
FT DOMAIN 296 308 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 309 332 7 (POTENTIAL).
FT DOMAIN 333 385 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 11 11 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 127 207 BY SIMILARITY.
FT LIPID 346 346 PALMITATE (POTENTIAL).
SQ SEQUENCE 385 AA; 43509 MW; 60A09B314E665B62 CRC64;

Query Match 19.7%; Score 440; DB 1; Length 385;
Best Local Similarity 25.5%; Pred. No. 4.2e-24;
Matches 119; Conservative 81; Mismatches 129; Indels 74; Gaps 16;

QY 23 BQETALVRLPLVVT-----PELPGRAKALVLTGLVIFA-----LALFGNAL 65
Db 19 ELFTKYLDR---YTPVSELALDPKELKDDSTLVEVQI--ILLFAYCSIIILGVIGNSL 74
QY 66 VFYVTVRSKAMRTVTNIFICSLALSDLLITFPFCIPVTMLQNISDNWLGGAFCIKMVPFVQ 125
Db 75 VIHVIKFKSNRTVTNFFIANLAVADLLVNTLCPLFTLVTLGGEWKLGPVLCHLVPAQ 134
QY 126 STAVVTILTWCTIAVERHOGVLVHPFKMKNQYTNRRAPFTMLGVVWLVAVIVGSPMHWVQ 185
Db 135 ALAVHVSVTTLTVIALDRHRCIVVHLESK--ISKRSIFLIIGVAVASALLASPLAIFRE 192
QY 186 ---LEIKYDFLYEKEHICCLLEWTSPP---VHOKLYTTFILVILFLPLWMLILYSKIGY 239
Db 193 YSLIEIIPDF---KIVVCEKWEPEGQLNGTIVSVSMLLIQVLPALIIISYATRI-- 246
QY 240 ELW--IKKRVGDGSLVLRTHGKMSKIARKKRAVIMMVTVVALFVCAVPFHVHMMIE 297
Db 247 --WTKLKNHVSFGAGNDHYH-----HRRQKTKMLVCVVVVFVSWLFPFHAPQLVSD 296
QY 298 YSNFE---KEYDDVTIKMIFAIVQIIGFNSICNPIYAFNMENFKKNVLSAVCVIYNK 354
Db 297 IDSQVLDLKEY-----KLIVTFVFIAMCSTFANPLLYGWNNNYRTAFITA----- 343
QY 355 TFSAPQRHGN-----SGITWNRKKAKEFLRENVPVEETKGEAFS 392
Db 344 -FQCEQLDLSIHPEVSAFAFKARKKLEAKKSQFP-----GDSFT 380

Search completed: October 28, 2003, 10:10:36
Job time : 26 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 10:08:35 ; Search time 100 Seconds
(without alignments)
1112.207 Million cell updates/sec

Title: US-10-070-241B-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRLRDHNL.....RHIALFRSELAENSPLDGSH 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mic:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvritus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1972	88.2	455	4 Q96P65	Q96P65 homo sapien
2	1603	71.7	365	11 Q8BHH0	Q8BHH0 mus musculus
3	493.5	22.1	375	13 Q57463	Q57463 brachydanio
4	485	21.7	417	11 Q924H0	Q924H0 mus musculus
5	485	21.7	417	11 Q8BKR6	Q8BKR6 mus musculus
6	484.5	21.7	374	13 Q9YHX1	Q9YHX1 gadus morhu
7	479.5	21.5	377	13 Q73733	Q73733 brachydanio
8	479	21.4	381	11 Q9ERC0	Q9ERC0 rattus norv
9	475.5	21.3	405	11 Q924N0	Q924N0 mus musculus
10	468.5	21.0	377	13 Q8QGM3	Q8QGM3 gallus gall
11	466.5	20.9	381	11 Q8BHW1	Q8BHW1 mus musculus
12	460.5	20.6	460	11 Q8BGL2	Q8BGL2 mus musculus
13	452	20.2	373	13 Q73734	Q73734 brachydanio
14	450.5	20.2	375	11 Q8BZF9	Q8BZF9 mus musculus
15	433.5	19.4	449	5 Q9VB87	Q9VB87 drosophila
16	431	19.3	385	13 Q8QFM1	Q8QFM1 gallus gall

17	424	19.0	432	11 Q924G9	Q924G9 rattus norv
18	422.5	18.9	425	4 Q9HBV6	Q9HBV6 homo sapien
19	422	18.9	365	13 Q8UVM7	Q8UVM7 lampetra fl
20	421.5	18.9	364	11 Q8BV78	Q8BV78 mus musculus
21	421.5	18.9	383	6 Q9GK75	Q9GK75 macaca mula
22	417	18.7	402	5 Q964E5	Q964E5 dugesia tig
23	410	18.3	372	11 Q942D4	Q942D4 cavia porce
24	408.5	18.3	375	6 Q97505	Q97505 sus scrofa
25	408	18.3	429	5 P92045	P92045 lymnaea sta
26	405.5	18.1	422	11 Q8VHD7	Q8VHD7 rattus norv
27	404.5	18.1	431	5 Q8T8D1	Q8T8D1 urechis uni
28	404	18.1	517	5 Q8T0S8	Q8T0S8 drosophila
29	403.5	18.1	402	5 Q20275	Q20275 caenorhabdi
30	401	17.9	519	5 Q9VAD2	Q9VAD2 drosophila
31	400.5	17.9	504	5 Q9VGX8	Q9VGX8 drosophila
32	400.5	17.9	600	5 Q9VM75	Q9VM75 drosophila
33	398	17.8	457	5 Q18534	Q18534 caenorhabdi
34	394.5	17.7	481	5 Q9VNM1	Q9VNM1 drosophila
35	394.5	17.7	485	5 Q8SZ35	Q8SZ35 drosophila
36	390.5	17.5	517	5 Q9VWR3	Q9VWR3 drosophila
37	390.5	17.5	678	5 Q94736	Q94736 stomoxys ca
38	390	17.4	465	5 Q44426	Q44426 lymnaea sta
39	387.5	17.3	540	5 Q9VRM0	Q9VRM0 drosophila
40	385	17.2	411	13 Q9W6I3	Q9W6I3 gallus gall
41	384.5	17.2	370	13 Q8UWL5	Q8UWL5 fugu rubrip
42	379	17.0	397	5 Q9NEA4	Q9NEA4 boophilus m
43	377.5	16.9	423	5 Q964D4	Q964D4 periplaneta
44	376.5	16.8	598	5 Q9VMQ9	Q9VMQ9 drosophila
45	375.5	16.8	311	11 Q8R1R6	Q8R1R6 cavia porce

ALIGNMENTS

RESULT 1

Q96P65	ID	Q96P65	PRELIMINARY;	PRT;	455 AA.
AC	Q96P65;				
DT	01-DEC-2001 (T-EMBLrel. 19, Created)				
DT	01-DEC-2001 (T-EMBLrel. 19, Last sequence update)				
DT	01-MAR-2003 (T-EMBLrel. 23, Last annotation update)				
DE	G protein-coupled receptor.				
GN	GPR103.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=21458557; PubMed=11574155;				
RA	Lee D.K., Nguyen T., Lynch K.R., Cheng R., Vanti W.B., Arkhitko O.,				
RA	Lewis T., Evans J.F., George S.R., O'Dowd B.F.;				
RT	"Discovery and mapping of ten novel G protein-coupled receptor				
RT	genes."				
RL	Gene 275:83-91(2001).				
DR	EMBL; AF411117; AAL26488.1; -				
DR	InterPro; HGNC:15565; GPR103.				
DR	InterPro; IPR000883; COX1.				
DR	InterPro; IPR000276; GPCR_Rhodpsn.				
DR	Pfam; PF00001; 7tm 1; 1.				
DR	PRINTS; PF00115; COX1.				
DR	PRINTS; PR00237; GPCRHOPOPSN.				
DR	PROSITE; PS00237; G PROTEIN RECP FL_1; 1.				
DR	PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.				
KN	Receptor.				
SQ	SEQUENCE. 455 AA; 51742 MW; AA9B541CCFD36742 CRC64;				

Query Match 88.2%; Score 1972; DB 4; Length 455;
Best Local Similarity 89.8%; Pred. No. 8.9e-159;
Matches 387; Conservative 0; Mismatches 6; Indels 38; Gaps 2;

Qy 1 MQALNITPEQFSRLRDHNLTRQFIALYRLVYTPPELPGRAKLALVLTGLFALAL 60
|||||


```

Db 314 NFOKELASLLSRC---RCWGPAESY-----ESFPLSTVSTGITKGSILNSGAS 359
QY 398 VKLCEQTEEEKKKR 412
Db 360 T---YQPHKKNLEQ 371

RESULT 4
Q924H0 PRELIMINARY; PRT; 417 AA.
AC Q924H0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuropeptide NPFF receptor.
GN GP274
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Liu Q., Guan X.-M., McDonald T.P., Jiang Q., Zeng Z., Marlene J.,
RA Williams D.L. Jr., Hong Y., Figueroa D., Clements M.K., Mallee J.,
RA Wang R., Evans J., Gould R., Austin C.P.;
RT "Identification and characterization of two cognate receptors for
RT mammalian FMRFamide-like neuropeptides."
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF330054; AAK94198.1; -.
DR MGD; MGI:1860130; Gpr74.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR InterPro; IPR005395; NPFF_receptor.
DR InterPro; IPR005397; NPFF_receptor2.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHOODPSN.
DR PRINTS; PR01570; NPFFRECEPTOR.
DR PRINTS; PR01572; NPFFRECEPTR2.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 417 AA; 47485 MW; AC0686514CDAF40C CRC64;

Query Match 21.7%; Score 485; DB 11; Length 417;
Best Local Similarity 31.0%; Pred. No. 7e-33;
Matches 121; Conservative 71; Mismatches 136; Indels 62; Gaps 11;

QY 17 DHNLTROFTALYRLRLPVLYTPELPGRAKLALVLTGVLIIFALALFGNALVYVYTRSKAM 76
Db 29 DINIT---YVNYLHQPV-----AAVFSSYLLIFVLCMVGNVVCVFIVRNRM 76
QY 77 RTVTNIFICSALSDLLITFCIPVTMLONTSDNWLGAFTCKMVPFVQSTAVVTEILTM 136
Db 77 HTVTNFFILNLAISDLVIGIFCMPTLLDNIAGWPGSSMCKISGLVQGISVAASVPTL 136
QY 137 TCIAVERHQLVHPFKMKWYTNRRAFITMLGVVLVAVIVGSP---MHWVQQLIKVDPL 193
Db 137 VAIAVDRPCVVPFKPK--LTVKTAFTVIIVINGLAITMTPSAIMLHVQ--BEKYRV 192
QY 194 YEKEH-----ICLEEWTSVPHQKIYTTFTILVILFLPLMVLILYSKIGYELWIKRV 247
Db 193 RLSSHNTSTVYWCREDWPRHEMRRIYTTVLFIATYIPLSLIVIMYARIGASLP----- 247
QY 248 GDGSLVLRTHGKMS---EMSKIARKKRAVIMVTVVVALFAVCWAPFHVHMMIEYSNPEKE 304
Db 248 ---KTAHACTGKQRPQVQHWVSKKQKVIKMLLTVALFILSWLPLMTLMMLSDYTDLSN 304
QY 305 YDDVTIKMIFAIVQIGFSNSTCNPIVYAFMNFENKKNVLSAVCYCIVNKTFSQAQRHGN 364
Db 305 KLRINIYIYFAHWAFLAFCSNVNPIIYGFNFENRFGQDAFOIC----- 350

Query Match 21.7%; Score 485; DB 11; Length 417;
Best Local Similarity 31.0%; Pred. No. 7e-33;
Matches 121; Conservative 71; Mismatches 136; Indels 62; Gaps 11;

QY 17 DHNLTROFTALYRLRLPVLYTPELPGRAKLALVLTGVLIIFALALFGNALVYVYTRSKAM 76
Db 29 DINIT---YVNYLHQPV-----AAVFSSYLLIFVLCMVGNVVCVFIVRNRM 76
QY 77 RTVTNIFICSALSDLLITFCIPVTMLONTSDNWLGAFTCKMVPFVQSTAVVTEILTM 136
Db 77 HTVTNFFILNLAISDLVIGIFCMPTLLDNIAGWPGSSMCKISGLVQGISVAASVPTL 136
QY 137 TCIAVERHQLVHPFKMKWYTNRRAFITMLGVVLVAVIVGSP---MHWVQQLIKVDPL 193
Db 137 VAIAVDRPCVVPFKPK--LTVKTAFTVIIVINGLAITMTPSAIMLHVQ--BEKYRV 192
QY 194 YEKEH-----ICLEEWTSVPHQKIYTTFTILVILFLPLMVLILYSKIGYELWIKRV 247
Db 193 RLSSHNTSTVYWCREDWPRHEMRRIYTTVLFIATYIPLSLIVIMYARIGASLP----- 247
QY 248 GDGSLVLRTHGKMS---EMSKIARKKRAVIMVTVVVALFAVCWAPFHVHMMIEYSNPEKE 304
Db 248 ---KTAHACTGKQRPQVQHWVSKKQKVIKMLLTVALFILSWLPLMTLMMLSDYTDLSN 304
QY 305 YDDVTIKMIFAIVQIGFSNSTCNPIVYAFMNFENKKNVLSAVCYCIVNKTFSQAQRHGN 364
Db 305 KLRINIYIYFAHWAFLAFCSNVNPIIYGFNFENRFGQDAFOIC----- 350

```

```

QY 365 SGITMRRKKAK----FSLR--ENPVEETKG 388
Db 351 -----QKXAKPOEAYSRLAKRNWINTSG 374

```

RESULT 5

```

Q8BXR6 PRELIMINARY; PRT; 417 AA.
AC Q8BXR6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuropeptide NPFF receptor.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=C57BL/6J; TISSUE=Body;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK050939; BAC34468.1; -.
SQ SEQUENCE 417 AA; 47449 MW; BB8D85E405D5786 CRC64;

Query Match 21.7%; Score 485; DB 11; Length 417;
Best Local Similarity 30.8%; Pred. No. 7e-33;
Matches 120; Conservative 73; Mismatches 135; Indels 62; Gaps 11;

QY 17 DHNLTROFTALYRLRLPVLYTPELPGRAKLALVLTGVLIIFALALFGNALVYVYTRSKAM 76
Db 29 DINIT---YVNYLHQPV-----AAVFSSYLLIFVLCMVGNVVCVFIVRNRM 76
QY 77 RTVTNIFICSALSDLLITFCIPVTMLONISDNWLGAFICKMVPFVQSTAVVTEILTM 136
Db 77 HTVTNFFILNLAISDLVIGIFCMPTLLDNIAGWPGSSMCKISGLVQGISVAASVPTL 136
QY 137 TCIAVERHQLVHPFKMKWYTNRRAFITMLGVVLVAVIVGSP---MHWVQQLIKVDPL 193
Db 137 VAIAVDRPCVVPFKPK--LTVKTAFTVIIVINGLAITMTPSAIMLHVQ--BEKYRV 192
QY 194 YEKEH-----ICLEEWTSVPHQKIYTTFTILVILFLPLMVLILYSKIGYELWIKRV 247
Db 193 RLSSHNTSTVYWCREDWPRHEMRRIYTTVLFIATYIPLSLIVIMYARIGASLP----- 247
QY 248 GDGSLVLRTHGKMS---KTIARKKRAVIMVTVVVALFAVCWAPFHVHMMIEYSNPEKE 304
Db 248 ---KTAHACTGKQRPQVQHWVSKKQKVIKMLLTVALFILSWLPLMTLMMLSDYTDLSN 304
QY 305 YDDVTIKMIFAIVQIGFSNSTCNPIVYAFMNFENKKNVLSAVCYCIVNKTFSQAQRHGN 364
Db 305 KLRINIYIYFAHWAFLAFCSNVNPIIYGFNFENRFGQDAFOIC----- 350

Query Match 21.7%; Score 485; DB 11; Length 417;
Best Local Similarity 30.8%; Pred. No. 7e-33;
Matches 120; Conservative 73; Mismatches 135; Indels 62; Gaps 11;

QY 17 DHNLTROFTALYRLRLPVLYTPELPGRAKLALVLTGVLIIFALALFGNALVYVYTRSKAM 76
Db 29 DINIT---YVNYLHQPV-----AAVFSSYLLIFVLCMVGNVVCVFIVRNRM 76
QY 77 RTVTNIFICSALSDLLITFCIPVTMLONISDNWLGAFICKMVPFVQSTAVVTEILTM 136
Db 77 HTVTNFFILNLAISDLVIGIFCMPTLLDNIAGWPGSSMCKISGLVQGISVAASVPTL 136
QY 137 TCIAVERHQLVHPFKMKWYTNRRAFITMLGVVLVAVIVGSP---MHWVQQLIKVDPL 193
Db 137 VAIAVDRPCVVPFKPK--LTVKTAFTVIIVINGLAITMTPSAIMLHVQ--BEKYRV 192
QY 194 YEKEH-----ICLEEWTSVPHQKIYTTFTILVILFLPLMVLILYSKIGYELWIKRV 247
Db 193 RLSSHNTSTVYWCREDWPRHEMRRIYTTVLFIATYIPLSLIVIMYARIGASLP----- 247
QY 248 GDGSLVLRTHGKMS---KTIARKKRAVIMVTVVVALFAVCWAPFHVHMMIEYSNPEKE 304
Db 248 ---KTAHACTGKQRPQVQHWVSKKQKVIKMLLTVALFILSWLPLMTLMMLSDYTDLSN 304
QY 305 YDDVTIKMIFAIVQIGFSNSTCNPIVYAFMNFENKKNVLSAVCYCIVNKTFSQAQRHGN 364
Db 305 KLRINIYIYFAHWAFLAFCSNVNPIIYGFNFENRFGQDAFOIC----- 350

```

RESULT 6

```

Q9YHX1 PRELIMINARY; PRT; 374 AA.
AC Q9YHX1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Neuropeptide Y-peptide YY receptor Yb.
GN NPYRB.
OS Gadus morhua (Atlantic cod).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Paracanthopterygii; Gadiformes; Gadidae; Gadus.
 OX NCBI_TaxID=8049;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99017378; PubMed=9802392;
 RA Arvidsson A.K., Wraith A., Jonsson-Rylander A.C., Larhammar D.;
 RT "Cloning of a neuropeptide Y/peptide YY receptor from the Atlantic
 RT cod: the Yb receptor";
 RL Regul. Pept. 75:39-43 (1998).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF073925; AAC02833.1; -;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 374 AA; 42260 MW; D2A9C9516C4998E2 CRC64;
 Query Match 21.7%; Score 484.5; DB 13; Length 374;
 Best Local Similarity 27.2%; Pred. No. 6.8e-33;
 Matches 108; Conservative 94; Mismatches 144; Indels 51; Gaps 10;
 QY 23 EQPI-----ALYRLPLVVTPELPGRKLA---LVLTGVLIPALALFGNALVYVVTSS 73
 Db 6 DQFINESHKPKANYSLCLAWDQECSPSSKSGITFLIVYSTMAIGVNSCLVFFVIAQ 65
 QY 74 KMRVTNIFISLALSDLLITPFCIPVTMLQNSDNWLGAPICKMPFVQSTAVVTBI 133
 Db 66 KEMHNVTNIFIANLSCDILMCFCLPVTLLIYTMDRWILGKALCKLTPFVQCISVTISI 125
 QY 134 LMTCTAVSRHGLVHPFKMKQYTNRRFTMLGVVWLVAVIVGSPMWH-----VQ 185
 Db 126 FSLVLAMERYQIILHP--TGWKPMPVQSTMAVGIIWVACLISVPFLSTVLDNLPLQN 183
 QY 186 LEIKYDFLYEKEHICCLEEWTSPVHOKIYTFILFLLPLMVMMLILSKIGYELWIKK 245
 Db 184 LSLPEP--GQDWLCTESWPTNSNRLAYTSLVQYFLPLGLIAACYLSIFLRLRRK 240
 QY 246 RVGDGSLVTIKGKNSKIARKKRAVIMVTVVVALFVCMAPFHVHVMIEYSNFEKY 305
 Db 241 -----DMVERARDSSRDNRKASRRINVMVGISTVALFVCMPLNINFTVFDW-----H 289
 QY 306 DQVTI---KMIPAVQIIIGFSNINCPVYAFMENKKNVLSAVCYCIVNKTSPAPR 361
 Db 290 HELMVSQNNLIPSVCHLVAMASTCVNVPVYIGFNSNFQKATLSHC---RCWGGAER 346
 QY 362 HGNSGITMRRKKAKFSLRENVPVEETKGEAFSDGNIEV 398
 Db 347 YENPLSAV-----STEVTKESHMSKGSISI 372
 RESULT 7
 QY7373 PRELIMINARY; PRT; 377 AA.
 AC O73733;
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Neuropeptide Y/peptide YY receptor Ya.
 GN NPYRYA OR NPYRYA.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98068842; PubMed=9407007;
 RA Lundell I., Berglund M.M., Starback P., Salaneck E., Gehlert D.R.,
 RA Larhammar D.;

*Cloning and characterization of a novel neuropeptide Y receptor
 subtypes in the zebrafish";
 RL DNA Cell Biol. 16:1357-1363 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98096393; PubMed=9434780;
 RA Ringvall M., Berglund M.M., Larhammar D.;
 RT Multiplicity of neuropeptide Y receptors: cloning of a third distinct
 RT subtype in the zebrafish";
 RL Biochem. Biophys. Res. Commun. 241:749-755 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99337783;
 RA Starback P., Lundell I., Fredriksson R., Berglund M.M., Yan Y.L.,
 RA Wraith A., Soderberg C., Postlethwait J.H., Larhammar D.;
 RT "Neuropeptide Y receptor subtype with unique properties cloned in the
 RT zebrafish: the zya receptor";
 RL Brain Res. Mol. Brain Res. 70:242-252 (1999).
 DR EMBL; AF037400; AAC41276.1; -;
 DR HSSP; P02699; 1F88.
 DR ZFIN; ZDB-GENE-980526-393; npyrya.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPS.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.
 KW Receptor.
 SQ SEQUENCE 377 AA; 42901 MW; 60562AD9C7BFE5D0 CRC64;
 Query Match 21.5%; Score 479.5; DB 13; Length 377;
 Best Local Similarity 29.7%; Pred. No. 1.8e-32;
 Matches 105; Conservative 83; Mismatches 131; Indels 35; Gaps 7;
 QY 46 LALVLTGVLIPALALFGNALVYVVTSSKAMRTVNIIFISLALSDLLITPFCIPVTMLQ 105
 Db 43 LTLVLCYCLNVLILGLLGNILICIMHQRPDPNPVTSLIANLSVSDILVSVFCLPFTVY 102
 QY 106 NISDNWLGAPICKMPFVQSTAVVTTELTCTAVSRHGLVHPFKMKQYTNRRFTM 165
 Db 103 TMDHWIGFALLCLMPFPVQCVSVTVSVLSVLIALERHQLILHP--SGWSPSVQAVIA 160
 QY 166 LGVWLVAVIVGSPM--WH-----VQOEIKYDFLYEKEHICCLEEWTSPVH 210
 Db 161 VLTWLLACVTSLSPLAFHLTSEBYSLSFPAPLSQLQV-----CLEVWPSQDH 208
 QY 211 QKIVTTFILVILFLLPLMVMMLILSKIGYELWIKKVGDSGLRTIKGKNSKIARKK 270
 Db 209 KIAYTTSLLLFQYCCFLMLLLCYLRIFLRQRRMLEROCSS--NEEDHRRVYMHKSR 266
 QY 271 AVIMVTVVVALFVCMAPFHVHVMIEYSNFEKYDDV--TIKMIPIAVQIIIGFSNINCP 329
 Db 267 INVMLATLVAFAVCMPLNANFV--ADCDQEVLPVCNHNHLSLCHLLAMSTCVNP 323
 QY 330 IYAFPMENFKKNVLSAVCYCIVNKTSPAPRHGNSGITMRRKKAKFSLRENVP 383
 Db 324 ILYGFLNSNFRKDVASVVLHCHFPLEDYSBHFPMSTMTNTDVSRTSFLRNSV 377
 RESULT 8
 QY9ERC0 PRELIMINARY; PRT; 381 AA.
 AC Q9ERC0;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Neuropeptide Y/peptide YY-Y2 receptor.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar; TISSUE=Small intestine;

```

RA Voisin T., Goumain M., LaBurthe M.;
RT "Cloning of a cDNA encoding a rat type 2 neuropeptide Y-peptide YY;"
RT receptor expressed in intestinal epithelial crypt cells.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AY004257; AAF89094.1;
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR001220; Lactin_legb.
DR Pfam; PF00001; 7tm1.1;
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS0307; LECTIN_LEGUMIN_BETA; 1.
KW Receptor.
SQ SEQUENCE 381 AA; 42510 MW; 6593227904B288BC7 CRC64;

Query Match 21.4%; Score 479; DB 11; Length 381;
Best Local Similarity 29.9%; Pred. No. 28-32;
Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;

QY 17 DHNLTRPFIALYRLPLVYTPLEFG-----RAKLAIVLTGVLIFALALFQGN 63
DQ 9 DENQTVKVELYSGPTTPRGELPPDEPELIDSTKLVEQVQLILAYCSIIILGVVGN 68
QY 64 ALVFVVTRESKAMRVTNIFICSIALSDLLITFCIPVTMLQNIQNSDNLGGAFICKMVPF 123
DQ 69 SLVHVIVIKFSMRTVTFNFFANLAVADLVNTLCPLPTLYTLTGEMKMGFVLCFLVEY 128
QY 124 VQSTAVVTEILTMTCIAVERHOGVLHPFKMKNQYTNRRFTMLGVVVLVAVVIGSPMHV 183
DQ 129 AGLAVQVSTILTIVIALDRHCIVYHLESK--ISKQISFLIIGLAWGSALASPLAIF 186
QY 184 QQ---LEIKYFLYKEHICCLBEW---TSPVHQKIYTTFILVILFLPLMVLMLYLSKI 237
DQ 187 REYSLEIIPDF---EIVACTEKPGBEKSYGVTVISLTLLIYVLPGLIISFYTRI 242
QY 238 GYELW--IKKRVGDSGLVLTTHGKMSKIARKKKKAVIMVTVVAFVAVCHAPFVHVHM 295
DQ 243 ----MSKLNHVSPGAASDHYH-----QRHKTKMLVGVVVVFAVSWLPLHAFQLA 290
QY 296 IYSNFE---KEYDDVTIKMFAIVQIIGFNSICNPIYAPMNNPKKNVLSAVCYCIV 352
DQ 291 VDISHVLDDKEY-----KLITVTHIAMCTFANPLLYGWNNSNYKAFUSA----- 339
QY 353 NKTSPAPQ---HNSGIMTKMKKAFSLRENPEVETKGEA 390
DQ 340 ---FRCEQLDAIHSEVSNFTKAKKNLEVKKNGNGLTDSFSEA 378

RESULT 9
Q924N0 PRELIMINARY; PRT; 405 AA.
AC Q924N0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE G-protein-coupled receptor 74.
GN GPR74.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Parker R.M., Herzog H.;
RT "Molecular cloning and characterisation of GPR74 a novel G-protein
RT coupled receptor closest related to the Y-receptor family.";
RL Brain Res. Mol. Brain Res. 0:0-0(2000).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; AF236084; AAK58514.1; -.
DR MGD; MGI:1860130; Gpr74.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR InterPro; IPR005395; NPFF_receptor.

```

```

DR InterPro; IPR005397; NPFF_receptor2.
DR Pfam; PF00001; 7tm1.1;
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PRINTS; PR01570; NPFFRECEPTOR.
DR PRINTS; PR01572; NPFFRECEPTR2.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
DR PROSITE; PS0307; LECTIN_LEGUMIN_BETA; 1.
KW G-protein coupled receptor; Receptor; Transmembrane.
SQ SEQUENCE 405 AA; 46248 MW; 8592068E1A17D64B CRC64;

Query Match 21.3%; Score 475.5; DB 11; Length 405;
Best Local Similarity 30.6%; Pred. No. 4.3e-32;
Matches 119; Conservative 74; Mismatches 135; Indels 61; Gaps 11;

QY 17 DHNLTRPFIALYRLPLVYTPLEFGRAKLALVLTGVLIFALALFQGNALVYVTVTRSKAM 76
DQ 32 DINIT---YVNYLHQPV-----AAVFISYLLIFVLCMVGNVTVGVFIVIRNRHM 79
QY 77 RTVNFIFICSIALSDLLITFCIPVTMLQNIQNSDNLGGAFICKMVPFVQSTAVVTEILTMT 136
DQ 80 HTVNFILNLALISDLLLVGIFCMETITLDNLIIAGNPGSSCKISGLVQGISVAASVFTL 139
QY 137 TCIAVERHOGVLHPFKMKNQYTNRRFTMLGVVVLVAVVIGSP---MHHVQOLEIKYDFL 193
DQ 140 VAIADVPRCVVYPPKPK--LTVKTAFTVTVIINGLAIAIMTPSAIMLHVQ--BEKYRV 195
QY 194 YEKSH-----ICLEEWTSVPVHOKIYTTFTILVILFLPLMVLMLYLSKIGVELWIKREV 247
DQ 196 RLSSHNTSTVYWCREDWPRHEMRRIYTVFAIYLAFLSLIVIMYARIGASLP----- 250
QY 248 GDGSLVLTTHGKE--MSKIARKKKRAVIMVTVVAFVAVCHAPFVHVHMIEYSNFEKEY 305
DQ 251 ---NTAAHCTQKRPVQCMYQKQKVIKMLLTVALFILSLWPLTWLMLSDYDLSNPK 307
QY 306 DDVTIKMFAIVQIIGFNSICNPIYAPMNNPKKNVLSAVCYCIVNKTSPAPQRHGNS 365
DQ 308 LRIINIVYPPFAHVLAFNCNSVNPFIYGFNFENRNGFQDAFQIC----- 352
QY 366 GITVMRKAK---PSLR--ENPVEETKG 388
DQ 353 -----QKAKPOEAYSURAKKNVINTSG 376

RESULT 10
Q8QGM3 PRELIMINARY; PRT; 377 AA.
AC Q8QGM3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Neuropeptide Y receptor 4.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Lundell I., Boswell T., Larhammar D.;
RT "Chicken Neuropeptide Y Family Receptor Y4; a Receptor with Equal
RT Affinity for Pancreatic Polypeptide, Neuropeptide Y and Peptide YY."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF410853; AAL84161.1; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1.1;
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
KW Receptor.
SQ SEQUENCE 377 AA; 42972 MW; C9EC6C00BDFD1F9E CRC64;

Query Match 21.0%; Score 468.5; DB 13; Length 377;
Best Local Similarity 30.6%; Pred. No. 1.6e-31;
Matches 102; Conservative 77; Mismatches 131; Indels 23; Gaps 6;

```

```

OY 58 LALFGNALVYVYVTRSKAMRTVTNIFICSIALSDLLITFFCIPVTMLQNSDNLGGAFI 117
Db 53 LGIVGNICLIAVIARQKKTNTNINILNLSIDLFMCVLCVLPFTVYVYVYVYVYVYVYVYV 112
OY 118 CKWVPVQSTAVVTELTWCTCAVEHGHGHPFKMKWQYTRRAFTMLGVVWLVAVIVG 177
Db 113 CKMTSFTQCTSVTSLSVLIALERHQHIIINP--TGMRPSTSQAYLGGVIGVITLACLMS 170
OY 178 SPMMHV-----OQLEIKYDFLYEKEHICLEEWTSVPHQKIYTTFFILVILFLPLMV 229
Db 171 LPFLTSLNDLYEGLSHIMNFSYDK--AICDSWSEQHRLIYTTLLILLOYCIPIFF 228
OY 230 MLILYSKIGYELWIKKRVGSGVLRTHGKESKIARKKKRAVIMVTVVAVFVCMWAPP 289
Db 229 IILCYLRIYLRQKQKDNPEKS-----EYSNRAVQLRRINILLASVAAFAVCWLP 280
OY 290 HVVHMWIEYSNEFEKEDVDVTIKMIFAIQIIGFSNSICNPVYAPMNEFKKNVLSAVCY 349
Db 281 HVFNITDWN--YKIISPCHHNLIIFSLCHLVAMASTVNPVYGFNGNFKKEVKSILIS 338
OY 350 CIVNKTFSQAORHNGSGI--TMMRKKAFFSLREN 381
Db 339 COHNSVTASIEDYDHLPLSTMTQTEVSKGSLSN 371

RESULT 11
OY Q8BWV1 PRELIMINARY; PRT; 381 AA.
AC Q8BWV1;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DE 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Neuropeptide Y receptor type 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Hippocampus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK049892; BAC33975.1; -.
DR EMBL; AK049892; BAC33975.1; -.
SQ SEQUENCE 381 AA; 42628 MW; 90787C051347884A CRC64;

```

```

Query Match 20.9%; Score 466.5; DB 11; Length 381;
Best Local Similarity 30.6%; Pred. No. 2.3e-31;
Matches 117; Conservative 71; Mismatches 121; Indels 73; Gaps 13;

OY 38 PELPGRAKLA----LVLTGVILFALFQNALVFVYVTRSKAMRTVTNIFICSIALSD 92
Db 38 PELIDSTKLVEQVILILAYCSIIILGVGNSLVHVLIKFSKMTVTNFFIANLAVDL 97
OY 93 LITFFCIPVTMLQNTSDNLGGAFICKWVPFVQSTAVVTEILTMTCAVERHGHGHPFK 152
Db 98 LVNTLCPLFTLTLYLMGKWKMPVLCHLVPAQGLAVQVSTITLTIVIALDRHRCIVYHLE 157
OY 153 MKWQYTRRAFTMLGVVWLVAVIVGSPMMHVQ---LEIKYDFLYEKEHICLEEW---T 206
Db 158 SK--ISKRSIFLIIGLAWGISALLASPLAIFREYSLIEIIPDF---EIVACTEKWPGEE 211
OY 207 SPVHOKIYTTFFILVILFLPLMVILYSKIGVELWIKKR--VGDGSLVLRTHGKEMSKI 264
Db 212 KSVYGTVSLSTLLILYVPLGLIISFSYTRI---WSKLNHVSGAASDHYH----- 260
OY 265 ARKKRAVIMVTVVAVFVCMWAPPVHVHMWIEYSNFE---KEYDDVTIKMIFAIQIIG 321
Db 261 -QRRHKWTKMLVCVVVFAVSMPLHAFQLAVDLSDHVLDAKEY-----KLIFTVFHIIA 314

```

```

OY 322 FSNSTCNPIVYAFMNEFKKNVLSAVCYCIVNKTFSQAOR-----HGNSGITMMRKKA 377
Db 315 MCSIFANPLLAGWNSNVRKAFLSA-----FRCEORLDAIHSEVSMTFKAKK----- 361
OY 378 LRENPFVEETKGEAFSDGNIEVK 399
Db 362 -----NLEVK 366

RESULT 12
OY Q8BG12 PRELIMINARY; PRT; 460 AA.
AC Q8BG12;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Orexin receptor type 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum, and Hypothalamus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573 (2002).
DR EMBL; AK038551; BAC30039.1; -.
DR EMBL; AK048781; BAC33457.1; -.
SQ SEQUENCE 460 AA; 52461 MW; D62A67C15BA67DCC CRC64;

```

```

Query Match 20.6%; Score 460.5; DB 11; Length 460;
Best Local Similarity 26.0%; Pred. No. 9.2e-31;
Matches 118; Conservative 97; Mismatches 164; Indels 75; Gaps 11;

OY 4 LNIITPEQFSRLLRDHLNLTREQFIAYLRPLRVYTPPLPGRAKLALVLTGVILFALFQ 63
Db 21 LNETQEPF--LNPTDYDDEFLR-YLMREYLH-----PKEYEWVLIAGIYVVFVALLGN 72
OY 64 ALVYVYVTRSKAMRTVTNIFICSIALSDLLITFFCIPVTMLQNSDNLGGAFICKWVPF 123
Db 73 VLVCVAVVAKNHHMRTVTNYFIVNLSDLVLTITCLPATLVVDITETWFFGOSLCKVIPY 132
OY 124 VOSTAVVTEILTMTCAVERHGHGHPFKMKWQYTRRAFTMLGVVWLVAVIVGSPMMHV 183
Db 133 LQTVSVSVSVLTSLCIALDRWYAICHPLMFK--STAKRARNISVIVITVSCIIIMIPQAI 190
OY 184 QOLEIKYDFLYEKEHI--CCLEEWTSVPHQKIYTTFFILVILFLPLMVILYSKIGVEL 241
Db 191 MECSSMLPGLANKTTFTVCDHGWGGEVYPMYHICPFLVYMAFLCLMILAYLQIFPKL 250
OY 242 WIKKRVGDSVLR-----THGKESKIARKKKRAVIMVTV 278
Db 251 WCRITPGTSTVYQKWKQKQVQSPQSGGQSKARISAVAAEIKQIRARRKTARMLMV-V 309
OY 279 VALFAVCWAPVHVHMWIE-YSNFEKEDVDVTIKMIFAIQIIGFSNSICNPVYAFNNE 337
Db 310 LLVFAICVLPISILUNLVKRVFGMTHETEDRETAVANWTFPSHNLVYANSAAPIIYNF 369
OY 338 NFKKNVLSAVCYCIVNKTFSQAORHNGSGITMMRKKAFFSLRENPFVEETKGEAFSDGN 397
Db 370 KFREEKAAFPSCCL-----GVHHRQCDRLARGR-- 397
OY 398 VKLCQTEEEKKLKRHLALFR--SELAENSPDGS 429
Db 398 ----TSTESRSLATQISNFDNVSKLSEHVVLTS 427

RESULT 13

```


073734	Q73734	PRELIMINARY;	PRT;	373 AA.
ID	Q73734	PRELIMINARY;	PRT;	373 AA.
AC	O73734;			
DT	01-AUG-1998 (TrEMBLrel. 07, Created)			
DT	01-AUG-1998 (TrEMBLrel. 07, Last sequence update)			
DE	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)			
DE	Neuropeptide Y/peptide YY receptor Yc.			
DE	NPYRYC OR NPYRYC.			
GN	NPYRYC			
OS	Brachydanio rerio (Zebrafish) (Danio rerio).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=98068842; PubMed=9407007;			
RX	Lundell I., Berglund M.M., Starback B., Salaneck E., Gehlert D.R.,			
RA	Larhammar D.;			
RA	"Cloning and characterization of a novel neuropeptide Y receptor			
RT	subtype in the zebrafish.";			
RT	DNA Cell Biol. 16:1357-1363(1997).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RP	MEDLINE=98096393; PubMed=9434780;			
RX	Ringvall M., Berglund M.M., Larhammar D.;			
RA	"Multiplicity of neuropeptide Y receptors: cloning of a third distinct			
RT	subtype in the zebrafish.";			
RT	Biochem. Biophys. Res. Commun. 241:749-755(1997).			
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).			
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.			
DR	EMBL; AF037401; AAC41277.1; -			
DR	ZFIN; ZDB-GENE-990415-175; npyryc.			
DR	InterPro; IPR000276; GPCR_Rhodopsn.			
DR	Pfam; PF00001; 7tm.1; 1.			
DR	PRINTS; PR00237; GPCRHHODPSN			
DR	PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.			
DR	PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.			
KW	G-protein coupled receptor; Receptor; Transmembrane.			
SK	SEQUENCE 373 AA; 41673 MW; 217DA7F51A940CFD CRC64;			
Query Match	20.2%; Score 452; DB 13; Length 373;			
Best Local Similarity	29.6%; Pred. No. 3.8e-30;			
Matches	96; Conservative 74; Mismatches 134; Indels 20; Gaps 6;			
QY	35 VYTPELFGRKALALVLTGLFALAFGLNALVFYVTRSKAMRTVNTFICSLASDLLI 94			
Db	22 VCPSPVSGTLLI VAYSTVI--AVGLVGNTCLVFIISRQKEMRVNTLILANLSCSDILM 79			
QY	95 TFCICPVTLQNLSDNLGGAFICKMVPFVOSTAVVTEILMTCTIAVERHQGLVHPFKMK 154			
Db	80 CVVCLPVTVTITLMDRVLTGELCKVTFPVQCMSTVSIFSLVILALERHQLIIHP--TG 137			
QY	155 WQYTNRRFTMLGVLVAVIVGSPMHVQQL-----EIKYDFLYEKEHICCLLEWTSP 208			
Db	138 WTPAAGHSYLAVATVMVACFISLPFLSFNLTNAPQNISLPNPFSDHVICHELWPE 197			
QY	209 VHOKIYTFILVILFLPLAVMLLILYSKIGVELYIKKVGVDGSLVURTHGEMSKIARKK 268			
Db	198 RNLRYTSTLLFOYCELELLILCYLIFLRLRRKDMVEQAT-----EARQKARGA 251			
QY	269 KRAVIMVTVVAFVACWAPPHVVMHLEYSNFEKEYDDVTIKMIFAIVQIIGFSNFCN 328			
Db	252 QRVNMLVIVIVAFLCWLPLNVNTTFDW--YHQLPACQHDVIFSCHLTATVASTCVN 309			
QY	329 PIVTAFMNEFKKNVLSAV--CYC 350			
Db	310 PVVYGFLNTNFQKELKATLQRCNC 333			
RESULT 14				
Q8BZF9				
ID	Q8BZF9			
AC	Q8BZF9;			
PRELIMINARY;	PRT;			375 AA.

RA Abril J.F., Agbayani A., An H.-J., Andrews-Sfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berzan B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriter A., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Corvelli J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hosain D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshtina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spiter E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Anantides P.G., Brandon R.C., Rogers Y.,
RA Banton J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn C., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003759; AAF5655.2;
DR FlyBase; Fggn0004842; Nery.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR PRINTS; PR00237; GPCR_Rhodopsn.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.

DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 449 AA; 51869 MW; FD8D0E3D70476EC1 CRC64;
Query Match 19.4%; Score 433.5; DB 5; Length 449;
Best Local Similarity 28.4%; Pred. No. 1.7e-28;
Matches 111; Conservative 75; Mismatches 156; Indels 49; Gaps 12;
QY 17 DHNLTREQFIALYRLRLPLVLTPELPGRKALALVLTGVLIFALALFGNALVYVYVTRSKAM 76
Db DYLLSEDMSAYFKIIVMLYIP-----IFIFALINGTVCYIYVSTPRM 121
QY 77 RTVTNIFICSIALSDLITFTFCIPVTMLQ-NISDNWLGAFICKMWPVQSTAVVTEILT 135
Db RTVTNYFIASLAIGDILMSPPCVSSFISLPIFNWPFGLALCHFNYSQAVSVLSAYT 181
QY 136 MTCIAVERHQGLVHPFKMKWYTNERRAFMLGVVLVAVIVGSPMWHVQOOLEIKYDFLYE 195
Db LVASIDRYIAIMWPLKPR--ITKRYATFIAGVWFIALATALPIPIVSGLDIPMSPMHT 239
QY 196 K-EHICCLLEWTSVPVHQKIYTFILVILFLPLMVLMLYLSKIGVELWIKKRVGDGSLR 254
Db KCEKVICREMPSPSTQEVYITLSLFALOFVVPVPLGVLTFTYATIRVWAKRPPGEAETNR 299
QY 255 TIHCKEMSKIARKKRAVIMMVTYVALFVAVCMAPFHVHMMIEYSNFEKEYDDVTIKMIF 314
Db D-----QRMARSKKRWKMKMLTVVIVFTCCWLPFNILQLLNDDEEF-AHWD--PLPYVW 350
QY 315 AIVOLIGFSNLCNPVIYVAPNNENFKKNVLSAV-----CYCI-----VNKTFS-- 357
Db 351 FAFHNLAMSHCCYPIIYCYNNARFRSGFVOLHMRMPLGLRWC-CLRSVGRDMNATSGT 409
QY 358 ---PAQRHNSG--ITWNRKKAKFSLRENPV 383
Db 410 PALPLNRMNTSTTVISARRKPRATSLRANPL 440

Search completed: October 28, 2003, 10:14:00
Job time : 106 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 28, 2003, 10:08:34 ; Search time 31 Seconds
(without alignments)
2328.243 Million cell updates/sec

Title: US-10-070-241B-1

Perfect score: 2235

Sequence: 1 MQALNITPEQFSRLRDHNL.....RHLLFRSELAENSLDLSGH 431

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2235	100.0	431	9	US-09-826-508-22
2	2235	100.0	431	9	US-09-899-532-2
3	2235	100.0	431	11	US-09-990-940-6
4	2232	99.9	431	12	US-10-272-983-38
5	2232	99.9	431	12	US-10-393-807-38
6	2135	95.5	415	12	US-10-080-263C-2
7	1914	85.5	433	11	US-09-990-940-16
8	1767	79.1	416	11	US-09-990-940-18
9	1268	56.7	245	15	US-10-225-567A-510
10	579	25.9	179	12	US-10-017-161-1164
11	502	22.5	420	10	US-09-866-248A-6
12	501.5	22.4	420	12	US-10-060-369-9
13	501.5	22.4	522	15	US-10-225-567A-512
14	479	21.4	381	12	US-10-188-619-4
15	479	21.4	381	15	US-10-067-649-5

16	474	21.2	381	12	US-10-188-619-6	Sequence 6, Appli
17	469.5	21.0	381	10	US-09-962-646-16	Sequence 16, Appl
18	469.5	21.0	381	10	US-09-292-973-5	Sequence 5, Appli
19	469.5	21.0	381	10	US-09-292-973-20	Sequence 20, Appl
20	469.5	21.0	381	12	US-10-188-619-2	Sequence 2, Appli
21	469.5	21.0	381	15	US-10-225-567A-201	Sequence 201, App
22	466.5	20.9	385	11	US-09-992-331-18	Sequence 18, Appl
23	466.5	20.9	385	16	US-10-262-313-18	Sequence 18, Appl
24	455	20.4	444	11	US-09-992-331-19	Sequence 19, Appl
25	455	20.4	444	12	US-10-060-369-11	Sequence 11, Appl
26	455	20.4	444	12	US-10-178-194-2	Sequence 2, Appli
27	455	20.4	444	15	US-10-225-567A-370	Sequence 370, App
28	455	20.4	444	16	US-10-262-313-19	Sequence 19, App
29	453	20.3	444	15	US-10-282-717-2	Sequence 2, Appli
30	440.5	19.7	345	12	US-10-359-285-5	Sequence 5, Appli
31	440	19.7	385	15	US-10-067-649-6	Sequence 6, Appli
32	437	19.6	432	10	US-09-866-248A-2	Sequence 2, Appli
33	434.5	19.4	347	12	US-10-254-905-9	Sequence 9, Appli
34	434.5	19.4	348	11	US-09-966-782A-9	Sequence 9, Appli
35	433.5	19.4	464	12	US-10-283-423-14	Sequence 14, Appl
36	433.5	19.4	464	12	US-10-213-821-14	Sequence 14, Appl
37	432.5	19.4	428	10	US-09-292-973-4	Sequence 4, Appli
38	432.5	19.4	430	10	US-09-866-248A-8	Sequence 8, Appli
39	432.5	19.4	430	15	US-10-225-567A-658	Sequence 658, App
40	431.5	19.3	345	12	US-10-254-905-10	Sequence 10, Appl
41	431.5	19.3	345	16	US-10-262-313-11	Sequence 11, Appl
42	431.5	19.3	346	8	US-08-899-112-32	Sequence 32, Appl
43	431.5	19.3	346	11	US-09-966-782A-10	Sequence 10, Appl
44	431.5	19.3	346	11	US-09-992-331-11	Sequence 11, Appl
45	431.5	19.3	346	11	US-09-771-287-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-09-826-508-22
; Sequence 22, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; TITLE OF INVENTION: and Polynucleotides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 22
; LENGTH: 431
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-826-508-22

Query Match	100.0%;	Score	2235;	DB	9;	Length	431;
Best Local Similarity	100.0%;	Pred. No.	1.7e-194;				
Matches	431;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MQALNITPEQFSRLRDHNLTREQFI	ALYRLPLVTPPELPGRAKALVLTGVLI	FALAL	60		
Db	1	MQALNITPEQFSRLRDHNLTREQFI	ALYRLPLVTPPELPGRAKALVLTGVLI	FALAL	60		
Qy	61	FGNALFYVVTSTRKARTVTNIFICS	LALSDLLITFFCIPVTMLQNISNWL	GGAFICKM	120		
Db	61	FGNALFYVVTSTRKARTVTNIFICS	LALSDLLITFFCIPVTMLQNISNWL	GGAFICKM	120		
Qy	121	VPFVQSTAVVTEILLTMTCTIAVER	HQGLVHPFKMKQVYTNRRFTMLGV	WLVAVIVGSPM	180		
Db	121	VPFVQSTAVVTEILLTMTCTIAVER	HQGLVHPFKMKQVYTNRRFTMLGV	WLVAVIVGSPM	180		
Qy	181	WHVQQLLEIKYDFLYEKEHICCLLE	EWTSVHQKIYTTFFILVILFLPLM	VMVLLYLSKIGYE	240		
Db	181	WHVQQLLEIKYDFLYEKEHICCLLE	EWTSVHQKIYTTFFILVILFLPLM	VMVLLYLSKIGYE	240		

Db 181 WHVQQLKDYFLYEKEHICCLLEWTSVPVHQKIYTTFFILVILFLPLVMVLLYLSKIGYE 240
QY 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
Db 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
Db 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
QY 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420
Db 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420
QY 421 LAENSPDLSGH 431
Db 421 LAENSPDLSGH 431

RESULT 2
US-09-899-532-2
; Sequence 2, Application US/09899532
; Patent No. US20020048791A1
; GENERAL INFORMATION:
; APPLICANT: Bloomquist, Brian T.
; APPLICANT: Zhelnin, Leonid
; TITLE OF INVENTION: Human Neuropeptide Y-Like G
; TITLE OF INVENTION: Protein-Coupled Receptor
; FILE REFERENCE: 02973.00040
; CURRENT APPLICATION NUMBER: US/09/899,532
; PRIOR FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/216,523
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-899-532-2

Query Match 100.0%; Score 2235; DB 9; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e-194;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLPALAL 60
Db 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLPALAL 60
QY 61 FGNALVFVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNLGGAFCIKM 120
Db 61 FGNALVFVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNLGGAFCIKM 120
QY 121 VPFVQSTAVVTEILMTCTIAVERHQGLVHPFKMKQWYTNRRAPTMVGWVWLVAIVGSPM 180
Db 121 VPFVQSTAVVTEILMTCTIAVERHQGLVHPFKMKQWYTNRRAPTMVGWVWLVAIVGSPM 180
QY 181 WHVQQLKDYFLYEKEHICCLLEWTSVPVHQKIYTTFFILVILFLPLVMVLLYLSKIGYE 240
Db 181 WHVQQLKDYFLYEKEHICCLLEWTSVPVHQKIYTTFFILVILFLPLVMVLLYLSKIGYE 240
QY 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
Db 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
Db 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
QY 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420
Db 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420

QY 421 LAENSPDLSGH 431
Db 421 LAENSPDLSGH 431
RESULT 3
US-09-990-940-6
; Sequence 6, Application US/09990940
; Publication No. US20030027252A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupta, Jamila S.
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030027252A1el Receptors
; FILE REFERENCE: 018781-007410US
; CURRENT APPLICATION NUMBER: US/09/990,940
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,841
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/257,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 431
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR346
US-09-990-940-6

Query Match 100.0%; Score 2235; DB 11; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.7e-194;
Matches 431; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLPALAL 60
Db 1 MQALNITPEQFSRLRDHNLTRQFIALYRLPLVYTPPELPGRAKLALVLTGVLPALAL 60
QY 61 FGNALVFVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNLGGAFCIKM 120
Db 61 FGNALVFVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNLGGAFCIKM 120
QY 121 VPFVQSTAVVTEILMTCTIAVERHQGLVHPFKMKQWYTNRRAPTMVGWVWLVAIVGSPM 180
Db 121 VPFVQSTAVVTEILMTCTIAVERHQGLVHPFKMKQWYTNRRAPTMVGWVWLVAIVGSPM 180
QY 181 WHVQQLKDYFLYEKEHICCLLEWTSVPVHQKIYTTFFILVILFLPLVMVLLYLSKIGYE 240
Db 181 WHVQQLKDYFLYEKEHICCLLEWTSVPVHQKIYTTFFILVILFLPLVMVLLYLSKIGYE 240
QY 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
Db 241 LMKRVGDSVLRTHGKEMSKIAKKKRAVIMVTVVAVCAWAPFHVHMMIEYSN 300
QY 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
Db 301 FEKEYDDVTIKMIFAIQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIWNKTFSPAQ 360
QY 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420
Db 361 RHNSGITMWRKAKFSLRENPEETKGEAFSDGNI EVKLCBOTBEKKLKHHLALFRSE 420

Db	301	FEKEYDDVTKMLPALVQIIIGFNSNCNPIVAFPMNFRKNLSAVCYICLVNKTFSQAQ	360
QY	361	RHNSGIGITMMRKAKFSLRENPNVEETKGEAFSDGNIIEVKLCEQTEEEKKUKRHLALFRSE	420
Db	361	RHNSGIGITMMRKAKFSLRENPNVEETKGEAFSDGNIIEVKLCEQTEEEKKUKRHLALFRSE	420
QY	421	LAENSPDLSGH	431
Db	421	LAENSPDLSGH	431
RESULT 5			
US-10-393-807-38			
; Sequence 38, Application US/10393807			
; Publication No. US20030175891A1			
; GENERAL INFORMATION:			
; APPLICANT: Chen, Ruoping			
; APPLICANT: Dang, Huang T.			
; APPLICANT: Liaw, Chen W.			
; APPLICANT: Lin, I-Lin			
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors			
; FILE REFERENCE: AREN0050			
; CURRENT APPLICATION NUMBER: US/10/393,807			
; CURRENT FILING DATE: 2003-03-21			
; PRIOR APPLICATION NUMBER: US/09/417,044			
; PRIOR FILING DATE: 1999-10-12			
; PRIOR APPLICATION NUMBER: 60/109,213			
; PRIOR FILING DATE: 1998-11-20			
; PRIOR APPLICATION NUMBER: 60/120,416			
; PRIOR FILING DATE: 1999-02-16			
; PRIOR APPLICATION NUMBER: 60/121,851			
; PRIOR FILING DATE: 1999-02-26			
; PRIOR APPLICATION NUMBER: 60/123,946			
; PRIOR FILING DATE: 1999-03-12			
; PRIOR APPLICATION NUMBER: 60/123,949			
; PRIOR FILING DATE: 1999-03-12			
; PRIOR APPLICATION NUMBER: 60/136,436			
; PRIOR FILING DATE: 1999-05-28			
; PRIOR APPLICATION NUMBER: 60/136,437			
; PRIOR FILING DATE: 1999-05-28			
; PRIOR APPLICATION NUMBER: 60/136,439			
; PRIOR FILING DATE: 1999-05-28			
; PRIOR APPLICATION NUMBER: 60/136,567			
; PRIOR FILING DATE: 1999-05-28			
; Remaining Prior application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 74			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 38			
; LENGTH: 431			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-393-807-38			
Query Match 99.9%; Score 2232; DB 12; Length 431;			
Best Local Similarity 99.8%; Pred. No. 3.2e-194;			
Matches 430; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
QY	1	MQALNITPEQFSRLLRDHNLTREQFIALYRLRPVYVTPPELPGRAKLAIVLTGVLIFALAL	60
Db	1	MQALNITPEQFSRLLRDHNLTREQFIALYRLRPVYVTPPELPGRAKLAIVLTGVLIFALAL	60
QY	61	FGNALVFYVYVTRSKAMKVTNIFCSIALSDLLITFFCIPVTMLQNI SDNLWGGAFCIKM	120
Db	61	FGNALVFYVYVTRSKAMKVTNIFCSIALSDLLITFFCIPVTMLQNI SDNLWGGAFCIKM	120
QY	121	VPFVQSTAVVTEMLTWTCAVERHQGLVHPFKMKQVYTNRRFTMLGWVWLVAIVVGS	180
Db	121	VPFVQSTAVVTEMLTWTCAVERHQGLVHPFKMKQVYTNRRFTMLGWVWLVAIVVGS	180
QY	181	WHVQOOLEIKYDFIYEKEHI CCLBEWTSPPVHQKIYTTFILVILFLLPLMWMILYISKIGYE	240
Db	181	WHVQOOLEIKYDFIYEKEHI CCLBEWTSPPVHQKIYTTFILVILFLLPLMWMILYISKIGYE	240

Db 421 LSENSTFGSGH 431

RESULT 8
US-09-990-940-18
; Sequence 18, Application US/09990940
; Publication No. US20030027252A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupte, Jamila S.
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030027252A1el Receptors
; FILE REFERENCE: 018781-007410US
; CURRENT APPLICATION NUMBER: US/09/990,940
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,841
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/257,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346b

US-09-990-940-18
Query Match 79.1%; Score 1767; DB 11; Length 416;
Best Local Similarity 80.1%; Pred. No. 4.9e-152;
Matches 339; Conservative 36; Mismatches 38; Indels 10; Gaps 3;

Qy	5	NITPEQSLRLDNLREQIALYRLPLVYTBELPGRKLAIVLTGVLIPALALFQNA	64
Db	4	NLTAEQSALLURLNLNLTQRAQFIHYGRPLVTPQLPARAKLALLVGNLIFALALFQNA	63
Qy	65	LVFVVTIRSKAMRTVTNIFICSLALSLLIIFFCIPVTMLQNI SDNMLGGAFICKMVPFV	124
Db	64	LVVIVTIRSKAMRTVTNIFICSLALSLLIIVFCIPVTMLQNVSDTMLGGAFICKMVPFV	123
Qy	125	QSTAVVTIELTMTCIATVERHOGLVHPFKQKQYNNRAFTMLGVVWLVAVTVGSPMHRVQ	184
Db	124	QCTAIVTEILTMTCIATVERHOGLVHPFKQKQYNNRAFTMLGVVWLVAVTVGSPMHRVQ	183
Qy	185	QLEIKYDPLYEKEHCCLLEWTSVPVHOKIYTTFTILVILFLPLMVLILYSKIGYELWIK	244
Db	184	KLEIKYDPLYEKEHCCLLEWSSPVHOKIYTTFTILVILFLPLMVLILYSKIGYELWIK	243
Qy	245	KRVGDSGLRTIHGKEMSKIARKKKRAVIMVTVVVALFAVCWAPPHVVMHMIYSNFEKE	304
Db	244	KRIGDGSVLRITIHGKEMFKIARKKKRAVIMVTVVVALFAVCWAPPHVVMHMIYSNFEKE	303
Qy	305	YDDVTIKMIFAIVQIIGFSNSICNPVIVAFNENPKKNVL SAVCYCIVNKTFTSPAQRHGN	364
Db	304	YDEVTKMIFAIVQIIGFSNSICNPVIVAFNENPKKNVL SAVCYCIVKETPPSARKHGS	363
Qy	365	SGITWRRKAKFSURENVEETKGEAFSDGNIEVKLCBOTEEKKKLKHLLALFRSELAEN	424
Db	364	SGAMVHRRAKLAARENVP-BIKGEAFGGSNIDIKWCQEPEKKKR-----RSKVA-S	413

Qy 425 SPL 427
Db 414 CPL 416

RESULT 9
US-10-225-567A-510
; Sequence 510, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 510
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-510

Query Match 56.7%; Score 1268; DB 15; Length 245;
Best Local Similarity 99.2%; Pred. No. 4.8e-107;
Matches 243; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy	186	LEIKYDPLYEKEHCCLLEWTSVPVHOKIYTTFTILVILFLPLMVLILYSKIGYELWIK	245
Db	1	MKLYDPLYEKEHCCLLEWTSVPVHOKIYTTFTILVILFLPLMVLILYSKIGYELWIK	60
Qy	246	RVGDSGLRTIHGKEMSKIARKKKRAVIMVTVVVALFAVCWAPPHVVMHMIYSNFEKEY	305
Db	61	RVGDSGLRTIHGKEMSKIARKKKRAVIMVTVVVALFAVCWAPPHVVMHMIYSNFEKEY	120
Qy	306	DDVTIKMIFAIVQIIGFSNSICNPVIVAFNENPKKNVL SAVCYCIVNKTFTSPAQRHGS	365
Db	121	DDVTIKMIFAIVQIIGFSNSICNPVIVAFNENPKKNVL SAVCYCIVNKTFTSPAQRHGS	180
Qy	366	GITWRRKAKFSURENVEETKGEAFSDGNIEVKLCBOTEEKKKLKHLLALFRSELAENS	425
Db	181	GITWRRKAKFSURENVEETKGEAFSDGNIEVKLCBOTEEKKKLKHLLALFRSELAENS	240
Qy	426	PLDSG 430	
Db	241	PLDSG 245	

RESULT 10
US-10-017-161-1164
; Sequence 1164, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: ABURATANI, HIROYUKI
; APPLICANT: AKIYAMA, YUTAKA
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; CURRENT FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 1164
; LENGTH: 179


```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-017-161-1164

Query Match      25.9%; Score 579; DB 12; Length 179;
Best Local Similarity 73.6%; Pred. No. 1.1e-44;
Matches 120; Conservative 10; Mismatches 13; Indels 20; Gaps 2;

QY 1 MQALNITPQFSLRLDHNLTREOFIALYRLRPLVYTPPELPGRAKALVLTGVLIFALAL 60
DQ 1 MQALNITPQFSLRLDHNLTREOFIALYRLRPLVYTPPELPGRAKALVLTGVLIFALAL 60
QY 61 FGNALVYVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNWLG----- 114
DQ 61 FGNALVYVYVTRSKAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNWLGKSRAPL 120
QY 115 --AFICKMVFF-----VQSTAVVTELTWTCTAVER 143
DQ 121 LRSFSLNPLPFRGSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLSLS 163

RESULT 11
US-09-866-248A-6
; Sequence 6, Application US/09866248A
; Publication No. US20020198367A1
; GENERAL INFORMATION:
; APPLICANT: Gerald, Christophe P.G.
; APPLICANT: Jones, Kenneth A.
; APPLICANT: Bonini, James A.
; APPLICANT: Borowsky, Beth
; TITLE OF INVENTION: DNA Encoding Mammalian Neuropeptide PF (NPFF) Receptors
; FILE REFERENCE: 1795/57155-A
; CURRENT APPLICATION NUMBER: US/09/866,248A
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 09/161,113
; PRIOR FILING DATE: 1998-09-25
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent In Ver. 2.0 - beta
; SEQ ID NO 6
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-248A-6

Query Match      22.5%; Score 502; DB 10; Length 420;
Best Local Similarity 30.1%; Pred. No. 3.2e-37;
Matches 128; Conservative 87; Mismatches 162; Indels 48; Gaps 12;

QY 14 LLRDHNLTRQOFIALYRLRPLVYTPPELPGRAKALVLTGVLIFALALFGNALVYVYVTR 73
DQ 26 LYSIDNIT---YVNYLHQPQV-----AAFIISYFLIFFLCWMGNTVVCFTVMRN 73
QY 74 KAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNWLGAFICKMVFFVQSTAVVTEI 133
DQ 74 KMHVTVNLFILNLAISDLLVGFICMPTLLDNIAGWFFGNTMCKISGLVQGISVAASV 133
QY 134 LTMCTIAVERHQGLVHPFKMKQYTNRAFTMLGVVVLVAVIVGSP---MHHVQQLLEIKY 190
DQ 134 FTLVAIAVDRFCQVYVFPKPK--LTIKTAFVIIMIWLVAITIMSPSAVMLHVQE-EKY 190
QY 191 DFLYEKEH-----ICCLEWTSVPHQKIYTTFTLVILFLLPLMVMMLIYSKIGYELWIKK 245
DQ 191 RVRLNSQNTSPVYWCREDWPNQENRKIYTVTLFANIYLAFLSLIVIMYGRIGISLF--- 247
QY 246 RVGDGSLVLRTHGKEMSK-----IARKKRAVIMVTVVVALFVCAWPAFHVHVMIEYSNF 301
DQ 248 ---RAAVPHT--GRKQVQWVSVSRKKQIKMLLIIVALLFILSLPLWTLMLLSYADL 302
QY 302 EKEYYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKNVLSAVCYCIVANKTSPAQ 361
DQ 303 SPNELQIINIYIYPPAHWLAFGNSSVNPFIYGFNENFRGRFQEAFLQLCQKRAKPM 362

US-10-070-241b-1.rapb
```

```
QY 362 HGNSGITMMRKAKKPSLRNPEVETKGEA-FSDGNIEVKLCQTEERKKLRHIALPRSE 420
DQ 363 YA-----LKAKSHVLINTSNQLVQESTFQPHGETLLYRKSAEKPOQE-----LVME 410
QY 421 LAENS 425
DQ 411 LKETT 415

RESULT 12
US-10-060-369-9
; Sequence 9, Application US/10060369
; Publication No. US20030139589A1
; GENERAL INFORMATION:
; APPLICANT: Zastawny, Roman
; TITLE OF INVENTION: G PROTEIN COUPLED RECEPTOR A4
; FILE REFERENCE: 2931-104
; CURRENT APPLICATION NUMBER: US/10/060,369
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 09/173565
; PRIOR FILING DATE: 1998-08-16
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patent In version 3.2
; SEQ ID NO 9
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Human
US-10-060-369-9

Query Match      22.4%; Score 501.5; DB 12; Length 420;
Best Local Similarity 30.1%; Pred. No. 3.5e-37;
Matches 124; Conservative 83; Mismatches 142; Indels 63; Gaps 12;

QY 14 LLRDHNLTRQOFIALYRLRPLVYTPPELPGRAKALVLTGVLIFALALFGNALVYVYVTR 73
DQ 26 LYSIDNIT---YVNYLHQPQV-----AAFIISYFLIFFLCWMGNTVVCFTVMRN 73
QY 74 KAMRTVTNIFICSALSDLLITFFCIPVTMLQNSDNWLGAFICKMVFFVQSTAVVTEI 133
DQ 74 KMHVTVNLFILNLAISDLLVGFICMPTLLDNIAGWFFGNTMCKISGLVQGISVAASV 133
QY 134 LTMCTIAVERHQGLVHPFKMKQYTNRAFTMLGVVVLVAVIVGSP---MHHVQQLLEIKY 190
DQ 134 FTLVAIAVDRFCQVYVFPKPK--LTIKTAFVIIMIWLVAITIMSPSAVMLHVQE-EKY 190
QY 191 DFLYEKEH-----ICCLEWTSVPHQKIYTTFTLVILFLLPLMVMMLIYSKIGYELWIKK 245
DQ 191 RVRLNSQNTSPVYWCREDWPNQENRKIYTVTLFANIYLAFLSLIVIMYGRIGISLF--- 247
QY 246 RVGDGSLVLRTHGKEMSK-----IARKKRAVIMVTVVVALFVCAWPAFHVHVMIEYSNF 301
DQ 248 ---RAAVPHT--GRKQVQWVSVSRKKQIKMLLIIVALLFILSLPLWTLMLLSYADL 302
QY 302 EKEYYDDVTIKMIFAIVQIIGFSNSICNPVYAFNENFKKN----- 342
DQ 303 SPNELQIINIYIYPPAHWLAFGNSSVNPFIYGFNENFRGRFQEAFLQLCQKRAKPM 362
QY 343 -VLSAVCYCIVANKTSPAQ-----HGNSGITMMRKAKKPSLRNPEVETK 387
DQ 363 YLAKASHVLINTSNQLVQESTFQPHGET--LLYRKSAEKPOQQLVMEELK 412

RESULT 13
US-10-225-567A-512
; Sequence 512, Application US/10225567A
; Publication No. US20030113798A1
; GENERAL INFORMATION:
; APPLICANT: LifeSpan Biosciences
; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenn C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTOR
; FILE REFERENCE: 1920-4-4
```

; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 512
; LENGTH: 522
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-225-567A-512

Query Match 22.4%; Score 501.5; DB 15; Length 522;
Best Local Similarity 30.1%; Pred. No. 4.6e-37;
Matches 124; Conservative 83; Mismatches 142; Indels 63; Gaps 12;

QY 14 LLDHNLTRQFIALYRLRLVYTPPELPGRAKLALVTGLVIFALFGNALFVYVTR 73
DB 128 LYSIDINIT--VYNYLHQPV-----AAIFIISYFLIFLCLMNGNTVVCFIVRN 175
QY 74 KAMRTVTNIFICSLALSDLIITFCIPVTMLQNSDNLGAFICKVVPFVQSTAVVTEI 133
DB 176 KMHVTVNLFILNALISDLIIVGLFCMPITLLDNLIAAGPFGNTWCKISGLVQGISVAASV 235
QY 134 LTWTCIAVERHQGLVHPFKMKQYTNRRFTMLGVWMLVAVIVGSP---MWHVQOLEIKY 190
DB 236 FTLVAIAVDFQCVVYFEK--LTIKTAFVIMIIIVLAINSPSAVMLHQE-EKY 292
QY 191 DFLYEKEH-----ICCLBEWTSVHQKIYTFILVILFLPLMWMLILYKIGVLMKK 245
DB 293 RVRLNSQNKTSPPVYWCREDPNQEMRKIYTVTLFANILYPLSLIVIMYGRIGISLP--- 349
QY 246 RVGDSVLRITGHKMSK---IARKKKRAVIMVMTVVALFVAVCWAPHVHVHMEYSNF 301
DB 350 ---RAAVPHT--GRNQEOMHVVRKKQKIIMLLIIVALLFILSWLPLNLMLSYADL 404
QY 302 EKEYDDVTIKMFAIVQIIGFSNICNPVIVAFNENFKKN----- 342
DB 405 SPNELQIINIVYFFAHWLAFGNSSVNPPIIYGFNENFRGFEAFQLOLCQKRAKPEA 464
QY 343 -VLSAVCYIVNKTFSQAOR-----HGNSGITMWRKAKFSLRENPNVEETK 387
DB 465 YTLKAKSHVLINTSNQLVQESTFQNPHGCT--LLYRKSAEKPOQELVMEELK 514

RESULT 14

US-10-188-619-4
; Sequence 4, Application US/10188619
; Publication No. US20030162944A1

GENERAL INFORMATION:

; APPLICANT: Christophe P.G. Gerald, et al.
; TITLE OF INVENTION: METHOD OF OBTAINING COMPOSITIONS COMPRISING Y2
; TITLE OF INVENTION: SPECIFIC COMPOUNDS

NUMBER OF SEQUENCES: 27

CORRESPONDENCE ADDRESS:

; ADDRESS: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/188,619

; FILING DATE:

CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/407,367

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 44742-AA-PCT-US/JPW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 381 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-10-188-619-4

Query Match 21.4%; Score 479; DB 12; Length 381;

Best Local Similarity 29.9%; Pred. No. 3.4e-35;
Matches 120; Conservative 78; Mismatches 144; Indels 60; Gaps 12;

QY 17 DHNLTRQFIALYRLRLVYTPPELPG-----RAKLALVTGLVIFALFGN 63
DB 9 DENQTVKVELYXSGPTTPRGELPPDEPELIDSKLVEVQVLLIAYCSIIILGVVGN 68
QY 64 ALVYVTVTRSKAMRTVTNIFICSLALSDLIITFCIPVTMLQNSDNLGAFICKVVPF 123
DB 69 SLVHVIVKFKSNRTVTNFIANLAVADLLVNTLCLPFTLTITLMGEWKNQGPVLCVLPY 128
QY 124 VQSTAVTEILTWTCIAVERHQGLVHPFKMKQYTNRRFTMLGVWMLVAVIVGSPMWHV 183
DB 129 AQGLAVQVSTITLTVIALDRHCIVVHLESK--ISKQISFLITGLANGVSALLASPLAIF 186
QY 184 QQ--LEIKYDFLYEKEHICCLEEM--TSPVHOKIYTTFILVILFLPLMWMLILYSKI 237
DB 187 REYSLIEITPDF---EIVACTEKNEGEKSVYGVTSLSLILLYVPLIGIISFSVTRI 242
QY 238 GYELW--IKERVGDGSLVLRITGHKMSKIAKKRAVIMVMTVVALFVAVCWAPHVHVHMM 295
DB 243 ---WSKLNHVSFGAASDHYH-----QRRHKTTKMLVCVVVFAVSWLPLHAFOLA 290
QY 296 IEYSNFE--KEYDDVTIKMFAIVQIIGFSNICNPVIVAFNENFKKNVLSAVCYIV 352
DB 291 VDISHVLDLKEY-----KLIFTVFHIIAMCSTFANPLLYGWNNSYRKAPLSA----- 339
QY 353 NKTFSQAOR---HGNSGITMWRKAKFSLRENPNVEETKGEA 390
DB 340 ---FRCEQLDAIHSEVSWTFFKKNLEVKKNGLDTSFSEA 378

RESULT 15

US-10-067-649-5

; Sequence 5, Application US/10067649

; Publication No. US20030100057A1

GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company

; TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY14, RELATED TO

; FILE REFERENCE: D0118 NP

; CURRENT APPLICATION NUMBER: US/10/067,649

; CURRENT FILING DATE: 2002-02-05

; PRIOR APPLICATION NUMBER: US 60/266,525

; PRIOR FILING DATE: 2001-02-05

; PRIOR APPLICATION NUMBER: US 60/329,897

; NUMBER OF SEQ ID NOS: 92

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 5

; LENGTH: 381

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-10-067-649-5

Query Match 21.4%; Score 479; DB 15; Length 381;

Best Local Similarity 29.9%; Pred. No. 3.4e-35;

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 07:53:00 ; Search time 654 Seconds
(without alignments)
10043.271 Million cell updates/sec

Title: US-10-070-241B-2
Perfect score: 2415
Sequence: 1 gccagagcgccaggaccc.....aaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1811591 seqs, 1359896290 residues

Total number of hits satisfying chosen parameters: 3623182

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US05_PUBCOMB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US04_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US03_PUBCOMB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US02_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US01_PUBCOMB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US00_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09D_PUBCOMB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1668.4	69.1	1710	9	US-09-899-532-3
2	1495	61.9	1495	12	US-10-080-263C-1
3	1296	53.7	1296	9	US-09-826-508-21
4	1296	53.7	1296	9	US-09-899-532-1
5	1296	53.7	1296	11	US-09-990-940-5
6	1294.4	53.6	1296	12	US-10-272-983-37
7	1294.4	53.6	1296	12	US-10-393-807-37
8	1078.4	44.7	1368	14	US-10-225-567A-509
9	994.6	41.2	1727	11	US-09-990-940-15
10	932	38.6	1772	11	US-09-990-940-17
11	542	22.4	940	12	US-10-017-161-1163
12	500	20.7	511	12	US-10-272-983-45
13	500	20.7	511	12	US-10-393-807-45
14	309	12.8	432	9	US-09-812-102-73
15	251.4	10.4	114793	12	US-10-148-806-3
16	249.8	10.3	31871	10	US-09-764-847-1403

17	249.8	10.3	31871	14	US-10-092-154-1403	Sequence 1403, Ap
18	249.6	10.3	288	12	US-10-229-058B-14	Sequence 14, Appl
19	249.2	10.3	684973	10	US-09-263-959-1	Sequence 1, Appl
20	248.4	10.3	250000	12	US-10-225-810-26	Sequence 26, Appl
21	248	10.3	1538	13	US-10-027-632-259313	Sequence 259313, Ap
c 22	247.8	10.3	31474	11	US-09-764-891-8149	Sequence 8149, Ap
c 23	247.6	10.3	300000	14	US-10-262-552-33	Sequence 33, Appl
24	247.4	10.2	11360	11	US-09-764-891-6508	Sequence 6508, Ap
25	247.4	10.2	11360	11	US-09-764-891-7415	Sequence 7415, Ap
26	246.8	10.2	1538	13	US-10-027-632-259314	Sequence 259314, Ap
27	246.8	10.2	1538	13	US-10-027-632-259315	Sequence 259315, Ap
c 28	246	10.2	203654	10	US-09-820-905-3	Sequence 3, Appl
c 29	245.6	10.2	1400	10	US-09-263-959-295	Sequence 295, Ap
c 30	245.6	10.2	35465	14	US-10-161-572-6	Sequence 6, Appl
c 31	245.6	10.2	36991	14	US-10-161-572-8	Sequence 8, Appl
c 32	245.6	10.2	684973	10	US-09-263-959-1	Sequence 1, Appl
33	245.6	10.2	1691139	14	US-10-067-514-1	Sequence 1, Appl
34	245.4	10.2	119596	14	US-10-270-336-3	Sequence 3, Appl
35	245.2	10.2	378361	11	US-09-901-136-3	Sequence 3, Appl
c 36	244.6	10.1	3512	13	US-10-027-632-114278	Sequence 114278, Ap
c 37	244.6	10.1	3512	13	US-10-027-632-114279	Sequence 114279, Ap
c 38	244.4	10.1	618	13	US-10-027-632-281878	Sequence 281878, Ap
39	244	10.1	1701	13	US-10-027-632-264344	Sequence 264344, Ap
40	244	10.1	1701	13	US-10-027-632-264345	Sequence 264345, Ap
41	244	10.1	1701	13	US-10-027-632-264346	Sequence 264346, Ap
42	244	10.1	1701	13	US-10-027-632-264347	Sequence 264347, Ap
43	244	10.1	1701	13	US-10-027-632-264348	Sequence 264348, Ap
c 44	243.8	10.1	301	10	US-09-860-670-183	Sequence 183, Ap
45	243.8	10.1	1843	13	US-10-027-632-100497	Sequence 100497, Ap

ALIGNMENTS

RESULT 1
US-09-899-532-3
; Sequence 3, Application US/09899532
; Patent No. US20020048791A1
; GENERAL INFORMATION:
; APPLICANT: Bloomquist, Brian T.
; TITLE OF INVENTION: Human Neuropeptide Y-Like G
; FILE OF INVENTION: Protein-Coupled Receptor
; FILE REFERENCE: 02973.00040
; CURRENT APPLICATION NUMBER: US/09/899,532
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/216,523
; PRIOR FILING DATE: 2000-07-06
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1710
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-899-532-3

Query Match	69.1%	Score 1668.4	DB 9	Length 1710
Best Local Similarity	99.9%	Pred. No. 0		
Matches 1669	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY	1	GCCAGAGCGCCAGACCCCTAGCGGTCCAGACCCAGACCCGCTGGCGGCGCTCG	60	
Db	41	GCCAGAGCGCCAGACCCCTAGCGGTCCAGACCCAGACCCGCTGGCGGCGCTCG	100	
QY	61	CTTAGGGAAGACGAGGAACCTTATTTGACCGGACATTTTTCCTACTGAG	120	
Db	101	CTTAGGGAAGACGAGGAACCTTATTTGACCGGACATTTTTCCTACTGAG	160	
QY	121	ATCAGTCTCCAGTCTTTCGCTTTCGCTCTTTATTCGTTGGTTCATCCCTGAGCTG	180	
Db	161	ATCAGTCTCCAGTCTTTCGCTTTCGCTCTTTATTCGTTGGTTCATCCCTGAGCTG	220	
QY	181	CTCTCTTTCCGAAACCTCCCGGGTGCAGCCTAGACCTCCCGCGGCTGACTCCAG	240	

```
Db 221 CTCTCTCTCCCGAACCTCCCGGGGTGACGCTTAGAGCCCTCCCGCGCGCTGACTCCAG 280
Qy 241 AGTAGAGGAGGAGCGCGCTCCGGCTGGTCCCCCGAAGCCCTGCTGCCCGCGAGATG 300
Db 281 AGTAGAGGAGGAGCGCGCTCCGGCTGGTCCCCCGAAGCCCTGCTGCCCGCGAGATG 340
Qy 301 CGGATGGCCAGCAGTAGCGGGCGGTGGCCCGCGTCCCGGAGCGCAGCAATGCGAG 360
Db 341 CGGATGGCCAGCAGTAGCGGGCGGTGGCCCGGTCCCGGAGCGCAGCAATGCGAG 400
Qy 361 CGCTTAACATTACCCCGAGAGTCTCTCGGCTGCTCGGGAACCAACCTGACGCGGG 420
Db 401 CGCTTAACATTACCCCGAGCAGTCTCTCGGCTGCTCGGGAACCAACCTGACGCGGG 460
Qy 421 AGCAGTTTCATCGCTGTAACCGGCTCGGACCGCTGCTCTACACCCAGAGCTGCCGGGAC 480
Db 461 AGCAGTTTCATCGCTGTAACCGGCTCGGACCGCTGCTCTACACCCAGAGCTGCCGGGAC 520
Qy 481 GGGCCAAAGCTGGCCCTCGTGCTCACCGGCGTGTCTCATCTTCGCCCTGGGCGCTCTTTGGCA 540
Db 521 GGGCCAAAGCTGGCCCTCGTGCTCACCGGCGTGTCTCATCTTCGCCCTGGGCGCTCTTTGGCA 580
Qy 541 ATGCTCTGCTGTCTACGTTGTGAACCGGAGCAAGCCCATGCGCACCGCTCAACCATCT 600
Db 581 ATGCTCTGCTGTCTACGTTGTGAACCGGAGCAAGCCCATGCGCACCGCTCAACCATCT 640
Qy 601 TTATCTGCTCCTTGGCGCTCAGTGACCTGCTCATCACCTTCTCTGTCATCCCGTCAACA 660
Db 641 TTATCTGCTCCTTGGCGCTCAGTGACCTGCTCATCACCTTCTCTGTCATCCCGTCAACA 700
Qy 661 TGTCTCAGAACATTTCCGACAACTGCTGGGGGTGCTTTCATTTGCAAGATGGTGCCAT 720
Db 701 TGTCTCAGAACATTTCCGACAACTGCTGGGGGTGCTTTCATTTGCAAGATGGTGCCAT 760
Qy 721 TTGTCAGTCTACCGTGTGTGACAGAAATCTCTAATGACCTGATGCTGTGTGGAAA 780
Db 761 TTGTCAGTCTACCGTGTGTGACAGAAATCTCTAATGACCTGATGCTGTGTGGAAA 820
Qy 781 GGCACACGAGCTGTGCAATCTTTTAAATGAAGTGGAATACACCAACCGAGGGCTT 840
Db 821 GGCACACGAGCTGTGCAATCTTTTAAATGAAGTGGAATACACCAACCGAGGGCTT 880
Qy 841 TCACATGCTAGTGTGTGCTGGCTGGTGCAGTCACTGATGATCACCCATGTGGCAG 900
Db 881 TCACATGCTAGTGTGTGCTGGCTGGTGCAGTCACTGATGATCACCCATGTGGCAG 940
Qy 901 TGCACAACTTGAGATCAATATGACTTCCTATATGAAAGGAACACATCTGCTGTAG 960
Db 941 TGCACAACTTGAGATCAATATGACTTCCTATATGAAAGGAACACATCTGCTGTAG 1000
Qy 961 AAGAGTGGACCAAGCCCTGTGCACAGAGATCTACACCACTTCATCCTTGTGATCCTCT 1020
Db 1001 AAGAGTGGACCAAGCCCTGTGCACAGAGATCTACACCACTTCATCCTTGTGATCCTCT 1060
Qy 1021 TCTCTCGCTCTTATGTTGATGCTTATCTGTACAGTAAATGTTGTTATGAACTTTGGA 1080
Db 1061 TCTCTCGCTCTTATGTTGATGCTTATCTGTACAGTAAATGTTGTTATGAACTTTGGA 1120
Qy 1081 TAAAGAAAAGAGTTGGGATGGTTAGTGTTCGAACTATTCATGAAAGAAATGTCGA 1140
Db 1121 TAAAGAAAAGAGTTGGGATGGTTAGTGTTCGAACTATTCATGAAAGAAATGTCGA 1180
Qy 1141 AAATAGCCAGGAAGAAACGAGCTGTCAATATGATGTTGAGTGGTGTCTCTTTG 1200
Db 1181 AAATAGCCAGGAAGAAACGAGCTGTCAATATGATGTTGAGTGGTGTCTCTTTG 1240
Qy 1201 CTGTGTGCTGGGACCAATTCATGTTGTCATATGATGATGATGATGATGATGATGATG 1260
Db 1241 CTGTGTGCTGGGACCAATTCATGTTGTCATATGATGATGATGATGATGATGATGATG 1300
Qy 1261 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
```

```
Db 1301 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1360
Qy 1321 CCACTCCATCTTAATCCCAATTCCTATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1361 CCACTCCATCTTAATCCCAATTCCTATGATGATGATGATGATGATGATGATGATGATG 1420
Qy 1381 TTTTGTCTGCAAGTTTGTATTGTCATAGTAATAAAACCTTCTCTCCAGCACAAGGCAATG 1440
Db 1421 TTTTGTCTGCAAGTTTGTATTGTCATAGTAATAAAACCTTCTCTCCAGCACAAGGCAATG 1480
Qy 1441 GAAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGTCTTCTCCAGAGAAATCCAG 1500
Db 1481 GAAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGTCTTCTCCAGAGAAATCCAG 1540
Qy 1501 TGGAGGAACCAAGGAGCAATTCAGTGATGCGCAACATTTGAGTCAAAATGTTGTAAC 1560
Db 1541 TGGAGGAACCAAGGAGCAATTCAGTGATGCGCAACATTTGAGTCAAAATGTTGTAAC 1600
Qy 1561 AGACAGAGGAGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGTCTGAATCTGGCTG 1620
Db 1601 AGACAGAGGAGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGTCTGAATCTGGCTG 1660
Qy 1621 AGAATTTCTCTTTAGACAGTGGGCAATTAATTAACAATACTTCAATAAT 1670
Db 1661 AGAATTTCTCTTTAGACAGTGGGCAATTAATTAACAATACTTCAATAAT 1710
```

RESULT 2

```
US-10-080-263C-1
; Sequence 1, Application US/10080263C
; Publication No. US20030143670A1
; GENERAL INFORMATION:
; APPLICANT: Bonini, James A.
; APPLICANT: Huang, Ling-Yan
; APPLICANT: Wilson, Amy
; TITLE OF INVENTION: DNA ENCODING SNORF44 RECEPTOR
; FILE REFERENCE: 1795/59370-A/JPW/ADM/ANX
; CURRENT APPLICATION NUMBER: US/10/080,263C
; CURRENT FILING DATE: 2002-02-20
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-080-263C-1
```

Query Match 61.9%; Score 1495; DB 12; Length 1495;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1495; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 181 CTCTCTTTCCCGAACCTCCCGGGTGCAGCTAGAGCCCTCCCGCGGCTGACTCCAG 240
Db 1 CTCTCTTTCCCGAACCTCCCGGGTGCAGCTAGAGCCCTCCCGCGGCTGACTCCAG 60
Qy 241 AGTAGAGGAAGGAGGCGGCTCCGGCTGGTCCCCCGAAGCCCTGCTGCCCGCAGATG 300
Db 61 AGTAGAGGAAGGAGGCGGCTCCGGCTGGTCCCCCGAAGCCCTGCTGCCCGCAGATG 120
Qy 301 CGGATGGCCAGCAGTAGCGGGCGGTGCGCGGTCCCGGAGCGGCAAGCAATGCGAG 360
Db 121 CGGATGGCCAGCAGTAGCGGGCGGTGCGCGGTCCCGGAGCGGCAAGCAATGCGAG 180
Qy 361 CGCTTAACATTACCCCGAGCAGTCTCTCGGCTGCTCGGGAACCAACCTGACGCGGG 420
Db 181 CGCTTAACATTACCCCGAGCAGTCTCTCGGCTGCTCGGGAACCAACCTGACGCGGG 240
Qy 421 AGCAGTTTCATCGCTGTAACCGGCTGCGACCGCTGCTACACCCAGAGTTCGCGGAC 480
Db 241 AGCAGTTTCATCGCTGTAACCGGCTGCGACCGCTGCTACACCCAGAGTTCGCGGAC 300
Qy 481 GGCACAACTGCTGCTGCTGCTACCGGCTGCTCATCTTCGCCCTGGCGCTCTTTGGCA 540
```

Db 301 GCGCAAGCTGGCCCTCGTGCTCACGGGGTGCTCACTTGGCCCTGGCGCTCTTTGGCA 360
Qy 541 ATGCTCTGGTGTCTACGTGTGACCCGACGAGGCGCATGGCAGCGCTCAACCAACTCT 600
Db 361 ATGCTCTGGTGTCTACGTGTGACCCGACGAGGCGCATGGCAGCGCTCAACCAACTCT 420
Qy 601 TTAATCTCTCTTGGCGCTGAGTGACCTGCTCATCACCTTCTTGTGCAATCCCGTCACCA 660
Db 421 TTAATCTCTCTTGGCGCTGAGTGACCTGCTCATCACCTTCTTGTGCAATCCCGTCACCA 480
Qy 661 TGCTCCAGAAATTTCCACAACTGGCTGGGGGTGCTTTCATTTGCAAGATGTCACAT 720
Db 481 TGCTCCAGAAATTTCCACAACTGGCTGGGGGTGCTTTCATTTGCAAGATGTCACAT 540
Qy 721 TTGTCCAGTCTACCGCTCTGTGACAGAAATCCTCACTATGACCTGCAATGCTGTGGAAA 780
Db 541 TTGTCCAGTCTACCGCTCTGTGACAGAAATCCTCACTATGACCTGCAATGCTGTGGAAA 600
Qy 781 GGCACACGGGACTTGTGATCCTTTTAAATGAAGTGGCAATACACCAACCGAAGGGCTT 840
Db 601 GGCACACGGGACTTGTGATCCTTTTAAATGAAGTGGCAATACACCAACCGAAGGGCTT 660
Qy 841 TCACAACTCTAGGTGTGCTGGCTGGTGGGAGTGCATGCTAGGATCACCCATGGGCACG 900
Db 661 TCACAACTCTAGGTGTGCTGGCTGGTGGGAGTGCATGCTAGGATCACCCATGGGCACG 720
Qy 901 TGCACAACTTGAGATCAAAATATGACTTCCATATGAAAAGGAACACATCTGCTCTTAG 960
Db 721 TGCACAACTTGAGATCAAAATATGACTTCCATATGAAAAGGAACACATCTGCTCTTAG 780
Qy 961 AAGAGTGAACCGACCTGTGCAACGAGAAGTCTACACCACTTCATCTTGTCTATCTCT 1020
Db 781 AAGAGTGAACCGACCTGTGCAACGAGAAGTCTACACCACTTCATCTTGTCTATCTCT 840
Qy 1021 TCCTCTCCCTCTATGTGTGATGCTTATCTGTACAGTAAATTTGTTATGAATTTGGA 1080
Db 841 TCCCTCCCTCTTATGTGTGATGCTTATCTGTACAGTAAATTTGTTATGAATTTGGA 900
Qy 1081 TAAAGAAAAGAGTTGGGGATGGTTTCAGTGTCTCGAACTTTCATGAGAAAAGAAATGTCCA 1140
Db 901 TAAAGAAAAGAGTTGGGGATGGTTTCAGTGTCTCGAACTTTCATGAGAAAAGAAATGTCCA 960
Qy 1141 AAATAGCCAGGAAGAAAGAGCTGTCTATATGATGGTGACAGTGGTGGCTCTCTTTG 1200
Db 961 AAATAGCCAGGAAGAAAGAGCTGTCTATATGATGGTGACAGTGGTGGCTCTCTTTG 1020
Qy 1201 CTGTGTGTGGGCACATTCATGTTGCTCATGATGATGATGATGATGATGATGATGATGAT 1260
Db 1021 CTGTGTGTGGGCACATTCATGTTGCTCATGATGATGATGATGATGATGATGATGATGAT 1080
Qy 1261 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
Db 1081 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Qy 1321 CCAACTCCATCTGTAATCCCATGCTATGCAATTTATGAATGAATTTCAAAAAAATG 1380
Db 1141 CCAACTCCATCTGTAATCCCATGCTATGCAATTTATGAATGAATTTCAAAAAAATG 1200
Qy 1381 TTTTCTGTGCACTTTGTTATGTCATAGTAAATATAAAACCTTCTCCAGCACAAAGGCATG 1440
Db 1201 TTTTCTGTGCACTTTGTTATGTCATAGTAAATATAAAACCTTCTCCAGCACAAAGGCATG 1260
Qy 1441 GAAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGTTTCCCTCAGAGAGAAATCCAG 1500
Db 1261 GAAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGTTTCCCTCAGAGAGAAATCCAG 1320
Qy 1501 TGGAGGAAACCAAAGAGAGCAATTCAGTGTGGCAACATGAAGTCAAAATGTTGTAAC 1560
Db 1321 TGGAGGAAACCAAAGAGAGCAATTCAGTGTGGCAACATGAAGTCAAAATGTTGTAAC 1380
Qy 1561 AGACAG 1620
Db 1381 AGACAG 1440

Qy 1621 AGAATTCCTCTTACAGTGGGCAATTAATTAACAATATCTTCAATAATTAATG 1675
Db 1441 AGAATTCCTCTTACAGTGGGCAATTAATTAACAATATCTTCAATAATTAATG 1495

RESULT 3

US-09-826-508-21
; Sequence 21, Application US/09826508
; Patent No. US20010025099A1
; GENERAL INFORMATION:
; APPLICANT: Nabil Elshourbagy
; APPLICANT: Lisa Vawter
; TITLE OF INVENTION: G Protein-Coupled Receptor Polypeptides
; FILE OF INVENTION: and Polynucleotides
; FILE REFERENCE: GP-70744USB
; CURRENT APPLICATION NUMBER: US/09/826,508
; CURRENT FILING DATE: 2001-04-05
; NUMBER OF SEQ ID NOS: 40
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 21
; LENGTH: 1296
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-826-508-21

Query Match 53.7%; Score 1296; DB 9; Length 1296;
Best Local Similarity 100.0%; Pred. No. 1.3e-285;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 354 ATGCAGGCGCTTAACATTAACCCGGAGCAGTTCTCTCGGTGCTGCGGACCAACCTG 413
Db 1 ATGCAGGCGCTTAACATTAACCCGGAGCAGTTCTCTCGGTGCTGCGGACCAACCTG 60
Qy 414 ACGGGGAGGAGTTCATCGCTCTGTCACGGCTGGAGCGCTCGTCTACACCCAGAGCTG 473
Db 61 ACGGGGAGGAGTTCATCGCTCTGTCACGGCTGGAGCGCTCGTCTACACCCAGAGCTG 120
Qy 474 CCGGAGCGCGCAAGCTGCTGCTCACCGGCGTGTCTCATCTTGGCCCTGGCGCTC 533
Db 121 CCGGAGCGCGCAAGCTGCTGCTCACCGGCGTGTCTCATCTTGGCCCTGGCGCTC 180
Qy 534 TTTGGCAATGCTCTGGTGTCTAGTGTGACCCGAGCAAGGCGCATGCGCACCGTACC 593
Db 181 TTTGGCAATGCTCTGGTGTCTAGTGTGACCCGAGCAAGGCGCATGCGCACCGTACC 240
Qy 594 AACATCTTTATCTGCTCTTGGCGCTCAGTGACCTGTCTCATCCTTCTTGTGATTTCC 653
Db 241 AACATCTTTATCTGCTCTTGGCGCTCAGTGACCTGTCTCATCCTTCTTGTGATTTCC 300
Qy 654 GTCACCATGCTCCAGAAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 713
Db 301 GTCACCATGCTCCAGAAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 360
Qy 714 GTGCCATTTGTCAGTGTACCGCTGTTGTGACAGAAATCTCATATGACCTGCAATGCT 773
Db 361 GTGCCATTTGTCAGTGTACCGCTGTTGTGACAGAAATCTCATATGACCTGCAATGCT 420
Qy 774 GTGGAAGGCAACCAAGGACCTGTGCATCTTTTAAATGAAGTGGCAATACACCAACCGA 833
Db 421 GTGGAAGGCAACCAAGGACCTGTGCATCTTTTAAATGAAGTGGCAATACACCAACCGA 480
Qy 834 AGGCTTTTCAATGCTAGGTGTGCTGCTGGTGGCAGTCACTGCTAGGATCAACCCATG 893
Db 481 AGGCTTTTCAATGCTAGGTGTGCTGCTGGTGGCAGTCACTGCTAGGATCAACCCATG 540
Qy 894 TGGCACGTGCAACCACTTGGAGTCAAAATATGACTTCTTATATGAAAAAGGAACACATCTGC 953
Db 541 TGGCACGTGCAACCACTTGGAGTCAAAATATGACTTCTTATATGAAAAAGGAACACATCTGC 600
Qy 954 TGCTTAGAGTGGACAGGCCCTGTGCAACGAGAGATCTACACCACTTCACTCTTGTGTC 1013
Db 601 TGCTTAGAGTGGACAGGCCCTGTGCAACGAGAGATCTACACCACTTCACTCTTGTGTC 660

[illegible]

US-10-225-567A-509

Query Match 44.7%; Score 1078.4; DB 14; Length 1368;
Best Local Similarity 91.3%; Pred. No. 6.2e-236;
Matches 1213; Conservative 0; Mismatches 1; Indels 114; Gaps 2;

QY 322 GCGGTGGCCCGCGTCCGCGGAGCGCACAGCAATCGAGCGGTAAACATTACCCCGGAGC 381
DB |||||
QY 155 GCGGTGGCCCGCGTCCGCGGAGCGCACAGCAATCGAGCGGTAAACATTACCCCGGAGC 214
DB |||||
QY 382 AGTTCTCTCGGCTGTCGCGGACCAACCTGACGCGGGAGAGTTCATCGCTCTGTACC 441
DB |||||
QY 215 AGTTCTCTCGGCTGTCGCGGACCAACCTGACGCGGGAGCAGTTCATCGCTCTGTACC 274
DB |||||
QY 442 GCGTGGACCGCTCGTACACCCAGAGCTGCGGGAGCGCCCAAGCTGGCCCTCGTGC 501
DB |||||
QY 275 GCGTGGACCGCTCGTACACCCAGAGCTGCGGGAGCGCCCAAGCTGGCCCTCGTGC 334
DB |||||
QY 502 TCACCGGCGTGTCTCTTTCGCCCTGGCGCTCTTTGGCAATGCTCTGGTGTCTACGTGG 561
DB |||||
QY 335 TCACCGGCGTGTCTCTTTCGCCCTGGCACCTTTTGGCAATGCTCTGGTGTCTACGTGG 394
DB |||||
QY 562 TGACCCGAGCAAGCGCATGCGCAGCGTCACCAACATCTTTATCTGCTTGGCGCTCA 621
DB |||||
QY 395 TGACCCGAGCAAGCGCATGCGCAGCGTCACCAACATCTTTATCTGCTTGGCGCTCA 454
DB |||||
QY 622 GTGACCTGCTATCACCTTCTTCTGCAITTCGCGTCCAGTGTCCAGAACATTTCCGACA 681
DB |||||
QY 455 GTGACCTGCTATCACCTTCTTCTGCAITTCGCGTCCAGTGTCCAGAACATTTCCGACA 514
DB |||||
QY 682 ACTGCTGGGGGTGCTTTCATTTGCAAGATGGTGCAITTTGTCAGTCTACCGCTGTGT 741
DB |||||
QY 515 ACTGCTGGGGGTGCTTTCATTTGCAAGATGGTGCCATTTGTCAGTCTACCGCTGTGT 574
DB |||||
QY 742 TGACAGAAATCCTCACTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 801
DB |||||
QY 575 TGACAGAAATCCTCACTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 634
DB |||||
QY 802 CTTTTAAATGAAGTGGCAATACCAACCGAGGGCTTTTCAATGCTAGTGTGTGTGT 861
DB |||||
QY 635 CTTTTAAATGAAGTGGCAATACCAACCGAGGGCTTTTCAATGCTAGTGTGTGTGT 694
DB |||||
QY 862 GCGTGGTGGAGTCTGCTAGATACCAATGTCAGCTGCAACAACTTGAGATCAAT 921
DB |||||
QY 695 GCGTGGTGGAGTCTGCTAGATACCAATGTCAGCTGCAACAACTTGAGATCAAT 754
DB |||||
QY 922 ATGACTTCTATGAAAGGAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 981
DB |||||
QY 755 ATGACTTCTATGAAAGGAACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
DB |||||
QY 982 ACCAGAGATCTACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
DB |||||
QY 815 ACCAGAGATCTACACCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 868
DB |||||
QY 1042 TCGTTATCTGACAGTAAATTTGTTATGAACTTTGTAAGAAAGAGTTGGGGATG 1101
DB |||||
QY 869 ----- 868
DB |||||
QY 1102 GTTCAGTGCTTCCAACTATTATGAAAGAAATGTCCAAATAGCCAGAGAGAGAAAC 1161
DB |||||
QY 869 -----GGAAGAGAGAAAC 880
DB |||||
QY 1162 GAGCTGCTAATTAAGATGAGTACAGTGGTGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCT 1221
DB |||||
QY 881 GAGCTGCTAATTAAGATGAGTACAGTGGTGGCTCTTTGCTGCTGCTGCTGCTGCTGCTGCT 940
DB |||||
QY 1222 ATGTTGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
DB |||||
QY 941 ATGTTGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1000
DB |||||
QY 1282 TCAAGATGATTTTGTATCGTGCAAAATTTATGGATTTTCCAACTCCATCTGTAATCCCA 1341
DB |||||
QY 1001 TCAAGATGATTTTGTATCGTGCAAAATTTATGGATTTTCCAACTCCATCTGTAATCCCA 1060
DB |||||

RESULT 9

US-09-990-940-15
; Sequence 15, Application US/09990940
; Publication No. US20030027252A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiagang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupta, Jamila S.
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030027252A1e1 Receptors
; FILE REFERENCE: 018781-007410US
; CURRENT APPLICATION NUMBER: US/09/990,940
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,841
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/257,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1737
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1302)
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346a
US-09-990-940-15

Query Match 41.2%; Score 994.6; DB 11; Length 1737;

Best Local Similarity 81.2%; Pred. No. 9.6e-217;

Matches 1172; Conservative 0; Mismatches 259; Indels 12; Gaps 1;

QY 354 ATGACGGCGCTTAACATTACCCCGGAGCAGTTCTCGGCTGCTGCGGACCAACCTG 413

DB 1 ATGACGGCGCTTAACATTACCCCGGAGCAGTTCTCGGCTGCTGCGGACCAACCTG 60

QY 414 ACGGGAGCAGTTCATCGCTCTGTACCGGTGCGACCGCTCGTCTACACCCAGAGCTG 473
DB 61 ACTCGGAAACAGTTCAATCGCTATCGGCTGCGACCGCTGCTACATCCGAGCTG 120
QY 474 CCGGGAACGCGCAAGCTTGGCCCTCGTGTCTACCGGCGTGTCTATCTTCGCGCTCGGCTC 533
DB 121 CCGCGCGCGCTAAACTGGCTTTGGCTGTGCTGAGCACTCAATTTTGGCTCGGCTC 180
QY 534 TTTGGCAATGCTCTGGTGTCTAGTGTGTGACCGCGCAGAGGCGCATCGCGACCGTCAAC 593
DB 181 TTTGGCAACTCTCTGGTCACTATGTGTGACCGCGCAGAGGCGCATCGCGACCGTCAAC 240
QY 594 AACATCTTTATCTCTCTCTGGCGCTCAGTCAACCTGTCTCATCACTCTCTCTGCAATCCC 653
DB 241 AACATCTTCTCTCTCTCTGGCACTCAGTCACTCTCTCTCTCTCTCTCTCTCTCTCT 300
QY 654 GTCACCATGCTCCAGAACATTTCCGACAACTGGGTGGGGGTGTTCATTTGCAAGATG 713
DB 301 GTCACGATGCTCCAGAACATCTCCGACAACTGGGTGGGTGGGTGGGTGGGTGGGTGG 360
QY 714 GTGCCATTTGTCAGTCTACCGCTGTGTGTGACAGAAATCCCTCACTAGACCTGCTTGC 773
DB 361 GTGCCCTTCTGTCAGTCTACCGCTGTGTGTGACAGAAATCCCTCACTAGTCTGCTGCT 420
QY 774 GTGAAAGGCAACCGAGGACTTGTGCTCTCTTTTAAATGAAGTGGCAATACACCAACCGA 833
DB 421 GTTGAGAGGCAACCGAGGACTCATCCATCTTTTAAATGAAGTGGCACTACCTACCGA 480
QY 834 AGGCTTTCAAACTAGT 893
DB 481 AGGCTTTCAAACTGT 540
QY 894 TGGCAGCTGCAACACTTGAGTCAAAATGACTTCTTATGATGAAATGAAGGAAACATCTGC 953
DB 541 TGGCAGCTGCAACCGCTCGAGTTAAGTATGACTTCTTATGAGAAAGAAACATGTCTGC 600
QY 954 TGCTTAGAAGTGTGACAGCCCTGTGCACCAAGAGATCTACACCACTTCACTCTCTGTC 1013
DB 601 TGTTTGAAGAGTGGCGCAGCCCATGACCAAGAAATCTACACCACTTCACTCTCTGTC 660
QY 1014 ATCT 1073
DB 661 ATCT 720
QY 1074 CTTTGGATAAAGAAAGTGTGGGATGTTTCAGTGTCTTCACTTATTCAGTGAAGAA 1133
DB 721 CTGTGGATCAAGAGAGAGTGTGGAGACAGTTCAGCACTTACAGCTATCCAGGGAAGAA 780
QY 1134 ATGTCCAAAATAGCAGGAAGAAACGAGTGTCTATATGATGTGTGACAGTGTGTGCT 1193
DB 781 ATGTCCAAAATAGCAGGAAGAAAGCGGCTGTCTGTTATGATGTGTGACAGTGTGTGCT 840
QY 1194 CTCCTTGT 1253
DB 841 CTCCTTGT 900
QY 1254 TTTGAAAAGGAATATGATGTGTCAATCAAGATGATTTTGTCTATCGTGCMAATATT 1313
DB 901 TTTGAAAAGGATATGATGTGTCAATCAAGATGATTTTGTCTATCGTGCMAAATATT 960
QY 1314 GGATTTTCCAACTCCATCTGTAATCCCAATGTGTCTATGATTTATGAATGAAATCTCAA 1373
DB 961 GGCTTTTCCAACTCCATCTGTAATCCCTTTGTGTATGATTTATGAATGAAATCTCAA 1020
QY 1374 AAAATGTTTGTCTGCAAGT 1433
DB 1021 AAGAAATTTTGTCTGCGGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
QY 1434 AGGCATGGAATTTCAAGGAATTAACAATGATGTGGGAAGAAAGCAAGTTTTCCTCTAGAG 1493
DB 1081 AAGCTTGAATTTCTGGGATTTCAATGATGTGAAAAGAGAGCAAGTTATCAGATCAAG 1140
QY 1494 AATCCAGTGGAGAAACCAAGAGAGAGCAATTCAGTGTGGAACATTTGAAGTCAATTTG 1553

DB 1141 CGTCCAGTGGCGAAGCCAAAGGAGACTTATTTCAGCGATGCCAAGTGTGATGTCARATTG 1200
QY 1554 TGTGAACAGACAGAGAGAGAAAGAAAGCTCAAAAGACATCTTGTCTCTTTAGTCTGAA 1613
DB 1201 TGTGAGCAGCAGGAGGAGAAAGCAACTCAAGCGACAGCTTGCCTTCTTTAGTCTGAA 1260
QY 1614 CTGGCTGAGAAATCTCTCTTTAGACAGTGGGCATTAATATAACAATATCTTCATAATTAA 1673
DB 1261 CTTTCTGAAACTCTACTTTCGCGAGTGGACATGACTGTAAATGATATCTCATAGCTAA 1320
QY 1674 TGCCTTTCAGATTGTAACCCAAAGAGAAAAATTTTGTAGCAAAAGTTCAAAATCTCTTTT 1733
DB 1321 TATCAITTTGAT-----GGAAAGTATTATTAGCAAAAGTTCAGGACTATTTTT 1368
QY 1734 AATCTTAAATGATGACAGAAAGAAACAAATCATGTTTCCATTAAAAATGACACAG 1793
DB 1369 TTTAAATGACAGAAAGAAACAAAGACATGTTTTCATTTTAAATGAACATAATACATAAC 1428
QY 1794 GCT 1796
DB 1429 ACT 1431

RESULT 10
US-09-990-940-17
; Sequence 17, Application US/09990940
; Publication No. US20030027252A1
; GENERAL INFORMATION:
; APPLICANT: Tian, Hui
; APPLICANT: Zhao, Jiayang
; APPLICANT: Chen, Jin-Long
; APPLICANT: Cutler, Gene
; APPLICANT: An, Songzhu
; APPLICANT: Dai, Kang
; APPLICANT: Gupte, Jamila S.
; APPLICANT: Tularik Inc.
; TITLE OF INVENTION: No. US20030027252A1el Receptors
; FILE REFERENCE: 018781-007410US
; CURRENT APPLICATION NUMBER: US/09/990,940
; CURRENT FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/252,841
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/257,636
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 60/261,377
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/279,554
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/280,696
; PRIOR FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 54
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 1772
; TYPE: DNA
; ORGANISM: Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1251)
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) TGR346b
US-09-990-940-17

Query Match 38.6%; Score 932; DB 11; Length 1772;
Best Local Similarity 82.0%; Pred. No. 1.9e-202;
Matches 1138; Conservative 0; Mismatches 235; Indels 15; Gaps 5;
QY 359 GGCGCTTAAACATTACCCGGAGCAGTTCTCTCGGCTGTCTCGGACCAACACCTGACGCG 418
DB 3 GTCTGTGAACCTTGACCCGGAGAGCTCTCGGCTGTCTCGGCTGTGACACCTGACGCG 62
QY 419 GGAGCAGTTCATCGCTCTGTACCGGCTGTGACCGCTGTCTGTACACCCAGAGCTGCCGG 478

Db 301 TCGTCTACACCCAGAGCTGCGGGAGCGGCGCAAGCTGGCCCTCGTGTCTACCGCGGTGC 360
Qy 514 TCATCTTCGCGCTGCGCTCTTTGGCAATGCTCTGGTGTCTACGTGTGAGACCCGACGA 573
Db 361 TCACTTCGCGCTGCGCTCTTTGGCAATGCTCTGGTGTCTACGTGTGAGACCCGACGA 420
Qy 574 AGGCCATGCGGACCGCTACCAACATCTTTATCTGCTCTCTTGGCGGTCTAGTGACCTGTCA 633
Db 421 AGGCCATGCGGACCGCTACCAACATCTTTATCTGCTCTCTTGGCGGTCTAGTGACCTGTCA 480
Qy 634 TCACCTCTCTGCAATCCCGTCCACATGCTCCAGACATTTCCGACAACTGGCTGGGG 693
Db 481 TCACCTCTCTGCAATCCCGTCCACATGCTCCAGACATTTCCGACAACTGGCTGGGG 540
Qy 694 GT 695
Db 541 GT 542

RESULT 12

US-10-272-983-45
; Sequence 45, Application US/10272983
; Publication No. US20030148450A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/272,983
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-272-983-45

Query Match 20.7%; Score 500; DB 12; Length 511;
Best Local Similarity 99.8%; Pred. No. 4.1e-104;
Matches 511; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 841 TCACATGCTAGTGTGCTGGCTGGCGAGTCATGTAGATCACCATGTGGCAGC 900
Db 1 TCACATGCTAGTGTGCTGGCTGGCGAGTCATGTAGATCACCATGTGGCAGC 60
Qy 901 TGCAACAACTTGAGATCAAAATATGACTTCTATATGAAAAAGAACACATCTGCTCTTAG 960
Db 61 TGCAACAACTTGAGATCAAAATATGACTTCTATATGAAAAAGAACACATCTGCTCTTAG 120

Qy 961 AAGAGTGGACAGCCCTGTGCACAGAGATCTACACACCTTCATCTCTGTCATCTCT 1020
Db 121 AAGAGTGGACAGCCCTGTGCACAGAGATCTACACACCTTCATCTCTGTCATCTCT 180
Qy 1021 TCCTTCCTGCTCTTATGATGATGCTTATTCGTACAGTAAATTTGGTTATGAATTTGGA 1080
Db 181 TCCTTCCTGCTCTTATGATGATGCTTATTCGTACAGTAAATTTGGTTATGAATTTGGA 239
Qy 1081 TAAAGAAAAGAGTTGGGATGTTTCAGTGTTCGAACTATTTCATGGAAGAAATGTCCA 1140
Db 240 TAAAGAAAAGAGTTGGGATGTTTCAGTGTTCGAACTATTTCATGGAAGAAATGTCCA 299
Qy 1141 AAATAGCCAGGAAGAAAGAGCTCTCATTTATGATGATGATGATGATGATGATGATG 1200
Db 300 AAATAGCCAGGAAGAAAGAGCTCTCATTTATGATGATGATGATGATGATGATGATG 359
Qy 1201 CTGTGTGCTGGGACCAATTCATGTTGTCATGATGATGATGATGATGATGATGATGATG 1260
Db 360 CTGTGTGCTGGGACCAATTCATGTTGTCATGATGATGATGATGATGATGATGATGATG 419
Qy 1261 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db 420 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
Qy 1321 CCAACTCCATCTGTAATCCCATTTGCTCTATGCA 1352
Db 480 CCAACTCCATCTGTAATCCCATTTGCTCTATGCA 511

RESULT 13

US-10-393-807-45
; Sequence 45, Application US/10393807
; Publication No. US20030175891A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Dang, Huong T.
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: Human Orphan G Protein Coupled Receptors
; FILE REFERENCE: AREN0050
; CURRENT APPLICATION NUMBER: US/10/393,807
; CURRENT FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: US/09/417,044
; PRIOR FILING DATE: 1999-10-12
; PRIOR APPLICATION NUMBER: 60/109,213
; PRIOR FILING DATE: 1998-11-20
; PRIOR APPLICATION NUMBER: 60/120,416
; PRIOR FILING DATE: 1999-02-16
; PRIOR APPLICATION NUMBER: 60/121,851
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/123,946
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/123,949
; PRIOR FILING DATE: 1999-03-12
; PRIOR APPLICATION NUMBER: 60/136,436
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,437
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,439
; PRIOR FILING DATE: 1999-05-28
; PRIOR APPLICATION NUMBER: 60/136,567
; PRIOR FILING DATE: 1999-05-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 45
; LENGTH: 511
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-393-807-45

Query Match 20.7%; Score 500; DB 12; Length 511;
Best Local Similarity 99.8%; Pred. No. 4.1e-104;


```
Qy 2073 CCAGCTACTTCGGAGCCTCAGGGGGAGAGATGGTGTGAACCGGGAGCGGAGCTTGCAG 2132
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23432 CCAGCTACTCGGGAGGCTGAGGCGAGGAGATGGCGGTGAACCGGGAGCGGAGCTTGCAG 23491
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2133 TGTCCGAGATCATGCCACTGCACTCCAGCCTGGGCGAAAGAGCGAGACTCCCGTCTCA 2192
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23492 TGAGCCAGATCGCGCCACCGCACTCCAGCCTGGGCGAGAGAGAGACTCCGTCTCAA 23551
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2193 AAAAAATTTTTTGAATAATTCGTAAACCAT 2225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23552 AAAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAAAAAACAAAAAAC 23584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: October 29, 2003, 12:05:06
Job time : 669 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 05:35:04 ; Search time 648 Seconds
(without alignments)
10060.406 Million cell updates/sec

Title: US-10-070-241B-2

Perfect score: 2415

Sequence: 1 gccagaggccaggaccct.....aaaaaaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_19Jun03.*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
- 25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2412	99.9	2415	22 AAF81819	Human G protein-co
2	2390.2	99.0	2448	25 AAS7631	Human G protein co
3	2389.6	98.9	2411	21 AAA23435	cDNA encoding huma
4	1668.4	69.1	1710	24 ABK14957	Human neurotrophide
5	1296	53.7	1296	22 AAH49502	Human GTP-binding
6	1296	53.7	1296	22 AAC87690	Human G protein-co
7	1296	53.7	1296	24 ABK6287	Human TGR346 cDNA.
8	1296	53.7	1296	24 ABK14956	Human neurotrophide

9	1294.4	53.6	1296	21 AAD01136	Human orphan G pro
10	1294.4	53.6	1296	21 AAA46037	Human G protein co
11	1293	53.5	1293	22 AAF81818	Human G protein-co
12	1289.6	53.4	1296	21 AAA46115	Human G protein co
13	1078.4	44.7	1368	25 AB242861	Human G protein-co
14	994.6	41.2	1737	24 ABR65292	Mouse TGR346a cDNA
15	932	38.6	1772	24 ABR65293	Mouse TGR346b cDNA
16	737	30.5	741	24 AAS98050	Human DNA for pote
17	500	20.7	511	21 AAD01143	Human orphan G pro
18	500	20.7	511	21 AAA46044	Human G protein co
19	309	12.8	432	22 AAI70302	Human GPCR-like pr
20	309	12.8	432	24 AB142350	Human galanin rece
21	309	12.8	432	24 AAS98092	Human DNA for pote
22	251.4	10.4	114793	22 AAD08215	Human genome from
23	251.2	10.4	506	22 ABA15628	Human nervous syst
24	251	10.4	3178	22 AAK80500	Human immune/haema
25	249.8	10.3	501	22 ABA15427	Human nervous syst
26	249.8	10.3	31871	23 ABK42516	Genomic sequence #
27	249.8	10.3	40742	22 AAK68089	Human immune/haema
28	249.8	10.3	40742	22 AAK79886	Human immune/haema
29	249.8	10.3	220895	24 ABK84798	Human cDNA differe
30	249.4	10.3	3178	22 AAK80499	Human immune/haema
31	249.2	10.3	267156	24 ABL68560	Kidney cancer rela
32	249	10.3	333	22 AAK66811	Human immune/haema
33	249	10.3	333	22 AAK66812	Human immune/haema
34	249	10.3	333	22 AAK66613	Human immune/haema
35	248.4	10.3	249999	25 AB280329	Human tramdorin ge
36	248	10.3	331	22 AAK78782	Human immune/haema
37	248	10.3	4124	22 AAS01154	5'-untranslated re
38	248	10.3	32145	22 AAK68491	Human immune/haema
39	248	10.3	32145	22 AAK68575	Human immune/haema
40	248	10.3	144460	21 AA293915	Olfactory receptor
41	247.8	10.3	31474	22 AAL05461	Human reproductive
42	247.8	10.3	31474	23 ABL98314	Human testicular a
43	247.4	10.2	11360	22 AAL03820	Human reproductive
44	247.4	10.2	11360	22 AAL04727	Human reproductive
45	247	10.2	17570	25 AB274059	Secreted protein g

ALIGNMENTS

RESULT 1

AAF81819

ID AAF81819 standard; cDNA; 2415 BP.

XX

AC AAF81819;

XX

DT 05-JUN-2001 (first entry)

XX

DE Human G protein-coupled receptor protein AQ27 encoding cDNA.

XX

Human; G protein-coupled receptor; AQ27; nootropic; neuroprotective;

hypotensive; antirheumatic; antiallergic; cardiant; antianaginal;

abortifacient; gene therapy; Alzheimer's disease; hypertension;

pregnancy termination; rheumatism; allergy; angina pectoris; ss.

XX

OS Homo sapiens.

XX

PH Key Location/Qualifiers

FT CDS 354..1649

FT /*tag= a

FT /product= "AQ27"

FT /note= "G protein coupled receptor"

XX

FN WO200116316-A1.

XX

PD 08-MAR-2001.

XX

PP 24-AUG-2000; 2000WO-JP05684.

XX

PR 27-AUG-1999; 99JP-0241530.

XX

QY 1741 AAGATGATGACAAAGAAACAAATCATGTTCCATTTAAATAATGACAGAGCTAGTC 1800
Db 1741 AAGATGATGACAAAGAAACAAATCATGTTCCATTTAAATAATGACAGAGCTAGTC 1800
QY 1801 CAAGTCAGTGTGTTTACAAACCAATNGATCAAAATCAATTAACANATTTCTGTGTCCT 1860
Db 1801 CAAGTCAGTGTGTTTACAAACCAATNGATCAAAATCAATTAACANATTTCTGTGTCCT 1860
QY 1861 TCTCATTTCCACCTGCTTCACTTGACTAGCCTTAAAAAGCAACATGGAAGCCAGGCACG 1920
Db 1861 TCTCATTTCCACCTGCTTCACTTGACTAGCCTTAAAAAGCAACATGGAAGCCAGGCACG 1920
QY 1921 GTGGCTCATGCTGTAAATCCAGACATTTTGGAGGCTTAGACGGGGGATCAAGAGTCA 1980
Db 1921 GTGGCTCATGCTGTAAATCCAGACATTTTGGAGGCTTAGACGGGGGATCAAGAGTCA 1980
QY 1981 GGAGATCAAAACCATCTGCTTAACACGCTGAAACCCCATCTCTGTAAAAATACAAAA 2040
Db 1981 GGAGATCAAAACCATCTGCTTAACACGCTGAAACCCCATCTCTGTAAAAATACAAAA 2040
QY 2041 TTAGCCGGGGGTGGTGGCGGACCTGTAGTCCAGCTACTTGGAGCCTCAGCGGGAG 2100
Db 2041 TTAGCCGGGGGTGGTGGCGGACCTGTAGTCCAGCTACTTGGAGCCTCAGCGGGAG 2100
QY 2101 AATGGTGTGAACCCGGGAGCGGAGCTTGCAATCCAGATCATGCCACTGCATCCA 2160
Db 2101 AATGGTGTGAACCCGGGAGCGGAGCTTGCAATCCAGATCATGCCACTGCATCCA 2160
QY 2161 GCCTGGCGGAAAGAGCGAGACTCCCGCTCTCAAAAAATTTTGTGAAAAATTCGTAA 2220
Db 2161 GCCTGGCGGAAAGAGCGAGACTCCCGCTCTCAAAAAATTTTGTGAAAAATTCGTAA 2220
QY 2221 CCATCTTTTAAGATTATTTTCACTGGATTTTAAAAATCTGTACAGAAATCAGGGTCT 2280
Db 2221 CCATCTTTTAAGATTATTTTCACTGGATTTTAAAAATCTGTACAGAAATCAGGGTCT 2280
QY 2281 TAGCTAGAGTTTCTCCACGAGTCACTGTAATGTGACTATGTAATCTAGATTGAA 2340
Db 2281 TAGCTAGAGTTTCTCCACGAGTCACTGTAATGTGACTATGTAATCTAGATTGAA 2340
QY 2341 TAAGAAAAATAAATAATATCTTCTTCTGAAAAAATAAATAAATAAATAAATAA 2400
Db 2341 TAAGAAAAATAAATAATATCTTCTTCTGAAAAAATAAATAAATAAATAAATAA 2400
QY 2401 AAAAAAAAAAAAAAGG 2415
Db 2401 AAAAAAAAAAAAAAGG 2415

RESULT 2
ABS57631
ID ABS57631 standard; cDNA; 2448 BP.
XX AC
XX AC
XX AC
XX ABS57631;
XX ABS57631;
DT 14-FEB-2003 (first entry)
DE Human G protein coupled receptor HGPBRMY2 cDNA SEQ ID 16.
XX
XX
XX Human; G-protein coupled receptor; HGPBRMY1; HGPBRMY2; immunosuppressive;
KW cardiant; neuroprotective; antiinflammatory; cytostatic; vulnery;
KW vaccine; gene therapy; autoimmune; cardiovascular; neural; reproductive;
KW haematopoietic; pulmonary; gastrointestinal; proliferation; cell cycle;
KW birth defect; aberrant phosphorylation; acute phase response; receptor;
KW signal transduction; hyperimmune activity; inflammatory; hypercongenital;
KW necrotic lesion; wound; organ transplant rejection; gene; disorder; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH 359..1654
FT CDS /*tag= a
FT /product= "HGPBRMY2"
FT

XX WO200268591-A2.
XX 06-SEP-2002.
XX 22-FEB-2002; 2002WO-US05281.
XX 23-FEB-2001; 2001US-270792P.
XX 23-FEB-2001; 2001US-270793P.
XX 06-JUN-2001; 2001US-296427P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Feder J, Ramanathan C, Nelson T, Mintier G, Cacace A, Barber L;
PI Kornacker M, Bol D;
XX WPI; 2003-058304/05.
XX P-PSDB; ABG73503.
XX New human HGPBRMY1 or HGPBRMY2 polynucleotide and polypeptide, useful
PT preventing, treating or ameliorating a disorder e.g., wound,
PT cardiovascular disorder or transplant rejection
XX Claim 1; Figure 6; 316pp; English.
XX This invention describes the novel human G-protein coupled receptors
CC (GPCR's), HGPBRMY1 or HGPBRMY2 which have immunosuppressive, cardiant,
CC neuroprotective, antiinflammatory, cytostatic and vulnery activity
CC and can be used in vaccines or for gene therapy. Pharmaceutical
CC compositions comprising HGPBRMY1 or HGPBRMY2 polypeptides or their
CC agonists or antagonists or modulators, or a HGPBRMY1- or
CC HGPBRMY2-specific antibody are useful for preventing, treating or
CC ameliorating a medical condition comprising autoimmune, cardiovascular,
CC neural, reproductive, haematopoietic, pulmonary, gastrointestinal or
CC proliferating disorder, a cell cycle or birth defect, a disorder related
CC to aberrant phosphorylation, acute phase responses or signal transduction
CC or to hyperimmune activity, an inflammatory or hypercongenital condition,
CC a necrotic lesion, a wound, organ transplant rejection or a condition
CC related to organ transplant rejection. This sequence encodes the human
CC HGPBRMY2 protein described in the disclosure of the invention.
XX
SQ Sequence 2448 BP; 665 A; 597 C; 575 G; 611 T; 0 other;
Query Match 99.0%; Score 2390.2; DB 25; Length 2448;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2392; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 7 GGGCCAGGACCTAGCTGGCGCTCCAGACCCAGACCGTGGCGGCGCTCGCCTTAG 66
Db 12 GGGCCAGGACCTAGCTGGCGCTCCAGACCCAGACCGTGGCGGCGCTCGCCTTAG 71
QY 67 GGAAGAGCAAGGAGAACTTTATTGACCGCGAACATTTTGGTCACTGAGATCGAG 126
Db 72 GGAAGAGCAAGGAGAACTTTATTGACCGCGAACATTTTGGTCACTGAGATCGAG 131
QY 127 TCTCCAGTGTCTTGGCTTCCCGCCTTTATCTGTTGGGTTTATCCTCAGTGTCTCC 186
Db 132 TCTCCAGTGTCTTGGCTTCCCGCCTTTATCTGTTGGGTTTATCCTCAGTGTCTCC 191
QY 187 TTTCCGAACTCTCCCGGGTGCAGCTAGAGCCTCCCGCGGCTAGTCCAGAGTAGA 246
Db 192 TTTCCGAACTCTCCCGGGTGCAGCTAGAGCCTCCCGCGGCTAGTCCAGAGTAGA 251
QY 247 GGAAGGAGCGGCTCCCGGCTGCTCCCGGAGCCTCGTCCCGGAGATGCGGATG 306
Db 252 GGAAGGAGCGGCTCCCGGCTGCTCCCGGAGCCTCGTCCCGGAGATGCGGATG 311
QY 307 GCCAGCCAGTAGCGGCGGCTGGCCCGCGTCCCGGAGCGCACAGCAATGCGGGCTTA 366
Db 312 GCCAGCCAGTAGCGGCGGCTGGCCCGCGTCCCGGAGCGCACAGCAATGCGGGCTTA 371
QY 367 ACATTACCCCGGAGCAGTTCTCTCGGCTGCTCGGAGACCAACCTGACCGGGAGCAGT 426

Db 372 ACATTACCCCGAGCAGTTCTCTCGGCTGCTGCGGACCAACAACCTGACGCGGAGCAGT 431
Qy
427 TCATCGCTCTCTGACCGGCTGCGACCGCTCGTCTACACCCCGAGAGCTGCGGAGCGGCA 486
Db
432 TCATCGCTCTCTGACCGGCTGCGACCGCTCGTCTACACCCCGAGAGCTGCGGAGCGGCA 491
Qy 487 AGCTGGGCCCTCGTGTCTACCGGCGTGTCTCATCTTTCGCCCTCGCGCTCTTTGGCAATGCTC 546
Db 492 AGCTGGGCCCTCGTGTCTACCGGCGTGTCTCATCTTTCGCCCTCGCGCTCTTTGGCAATGCTC 551
Qy 547 TGGTGTCTCTAGTGGTGTGACCGGCGAGCGAGGCGCATGCGCACCGTCAACCAATCTTTATCT 606
Db 552 TGGTGTCTCTAGTGGTGTGACCGGCGAGCGAGGCGCATGCGCACCGTCAACCAATCTTTATCT 611
Qy 607 GCTCCTTGGCGCTCAGTGACCTGTCTCATCACTTCTTCTGCAATCCCGTCAACCATGCTCC 666
Db 612 GCTCCTTGGCGCTCAGTGACCTGTCTCATCACTTCTTCTGCAATCCCGTCAACCATGCTCC 671
Qy 667 AGAACATTCGACMACTGGCTGGGGGTGCTTTCATTTGGCAAGATGTCATTTGTC 726
Db 672 AGAACATTCGACMACTGGCTGGGGGTGCTTTCATTTGGCAAGATGTCATTTGTC 731
Qy 727 AGTCTACCGCTGTGTGACGAATCCCTCACTATGACCTGCAATGCTGTGGAAGGCACC 786
Db 732 AGTCTACCGCTGTGTGACGAATCCCTCACTATGACCTGCAATGCTGTGGAAGGCACC 791
Qy 787 AGGAATGTGCACTCTTTTAAATGAAGTGGCAATACACCAACCGAAGGCTTTCACAA 846
Db 792 AGGAATGTGCACTCTTTTAAATGAAGTGGCAATACACCAACCGAAGGCTTTCACAA 851
Qy 847 TGCTAGTGTGCTGCTGGCTGGCAGTCATGCTAGGATCACCATGTCGCACTGCAAC 906
Db 852 TGCTAGTGTGCTGCTGGCTGGCAGTCATGCTAGGATCACCATGTCGCACTGCAAC 911
Qy 907 AACTTGAGATCAATATGACTTCTTATGAAAGGAACACATCTGCTGTAGAGAGT 966
Db 912 AACTTGAGATCAATATGACTTCTTATGAAAGGAACACATCTGCTGTAGAGAGT 971
Qy 967 GGACAGCCCTGTGCAACAGAGATCTACACACCTTCATCCTTGTCACTCTTCTCTCC 1026
Db 972 GGACAGCCCTGTGCAACAGAGATCTACACACCTTCATCCTTGTCACTCTTCTCTCC 1031
Qy 1027 TGCCTCTTATCGTGATCTTATCTGTACAGTAAATTTGGTATGAACTTTGGTAAAGA 1086
Db 1032 TGCCTCTTATCGTGATCTTATCTGTACAGTAAATTTGGTATGAACTTTGGTAAAGA 1091
Qy 1087 AAAGAGTTGGGATGGTTCAGTGTCTCGAATCTTATGAAAGGAAGATGTCGAAATAG 1146
Db 1092 AAAGAGTTGGGATGGTTCAGTGTCTCGAATCTTATGAAAGGAAGATGTCGAAATAG 1151
Qy 1147 CCAGAGAGAGAAACGAGCTGTCAATATGATGGTGACAGTGGTGGCTCTCTTGTGTGT 1206
Db 1152 CCAGAGAGAGAAACGAGCTGTCAATATGATGGTGACAGTGGTGGCTCTCTTGTGTGT 1211
Qy 1207 GCTGGGACCACTTCATCTGTCATATGATGATGATGATGATGATGATGATGATGATGAT 1266
Db 1212 GCTGGGACCACTTCATCTGTCATATGATGATGATGATGATGATGATGATGATGATGAT 1271
Qy 1267 ATGATGATGTCACAAATCAAGATGATTTTGTGATGTCGTCGTCGTCGTCGTCGTCGTC 1326
Db 1272 ATGATGATGTCACAAATCAAGATGATTTTGTGATGTCGTCGTCGTCGTCGTCGTCGTC 1331
Qy 1327 CCATCTGTAATCCCATTTGCTATGCAATTTATGATGATGATGATGATGATGATGATGAT 1386
Db 1332 CCATCTGTAATCCCATTTGCTATGCAATTTATGATGATGATGATGATGATGATGATGAT 1391
Qy 1387 CTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
Db 1392 CTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1451
Qy 1447 CAGGAATTTACAAATGATGCGGAGAGAAACCAAGTCTTCCCTCAGAGAGATCCAGTGGAGG 1506
Db 1452 CAGGAATTTACAAATGATGCGGAGAGAAACCAAGTCTTCCCTCAGAGAGATCCAGTGGAGG 1511

Qy 1507 AACCAAGAGAGAGCATTTCAGTGATGGCAACATTTGAAGTCAAAATTCGTGAAACAGACAG 1566
Db 1512 AACCAAGAGAGAGCATTTCAGTGATGGCAACATTTGAAGTCAAAATTCGTGAAACAGACAG 1571
Qy 1567 AGGAGAGAGAGAGAGCTCAAAACGACATCTTGTCTCTCTTTAGGTCTGAACTGGCTGAGAAAT 1626
Db 1572 AGGAGAGAGAGAGAGCTCAAAACGACATCTTGTCTCTCTTTAGGTCTGAACTGGCTGAGAAAT 1631
Qy 1627 CTCCTTTAGACAGTGGGCATTATTAACAATATCTTCAATATTAATTAATGCTTCAAT 1686
Db 1632 CTCCTTTAGACAGTGGGCATTATTAACAATATCTTCAATATTAATTAATGCTTCAAT 1691
Qy 1687 GTAACCCAAAGAGAGAAATTTTGAACAAAGTCAAAATCTCTTTAAATCTTTAAAGATG 1746
Db 1692 GTAACCCAAAGAGAGAAATTTTGAACAAAGTCAAAATCTCTTTAAATCTTTAAAGATG 1751
Qy 1747 ATGACAAAGAGAGAGAAACAAATCATCTTTTCCATTTAAATAACACAGAGCTAGTCCAAAGT 1806
Db 1752 ATGACAAAGAGAGAGAAACAAATCATCTTTTCCATTTAAATAACACAGAGCTAGTCCAAAGT 1811
Qy 1807 CAGTGATGTTTCAACCAATNGATCAATCAATTTAACANATTTCTGTGTTCTCTTCAT 1866
Db 1812 CAGTGATGTTTCAACCAATNGATCAATCAATTTAACANATTTCTGTGTTCTCTTCAT 1871
Qy 1867 TCCCACTGCTTCACTTTGACTAGCTTTAAAAAAGCAACATGGAAGGCGCAGCAGCGTGGCT 1926
Db 1872 TCCCACTGCTTCACTTTGACTAGCTTTAAAAAAGCAACATGGAAGGCGCAGCAGCGTGGCT 1931
Qy 1927 CATGCCCTGTAATCCACAGCCTTTGGAGGCGCTAGACGGGCGGATCAGAGGTGAGAGAT 1986
Db 1932 CATGCCCTGTAATCCACAGCCTTTGGAGGCGCTAGACGGGCGGATCAGAGGTGAGAGAT 1991
Qy 1987 CAACCAATCCTGGCTAAACACGCTGAAACCCCATCTCTGCTAAATAATACAAAATTTAGCC 2046
Db 1992 CAACCAATCCTGGCTAAACACGCTGAAACCCCATCTCTGCTAAATAATACAAAATTTAGCC 2051
Qy 2047 GGGCGTGGTGGCGGACCTGTAGTCCAGCTACTTTGGAGGCTCAGGCGGAGAAATGGT 2106
Db 2052 GGGCGTGGTGGCGGACCTGTAGTCCAGCTACTTTGGAGGCTCAGGCGGAGAAATGGT 2111
Qy 2107 GTAAACCCGGAGGCGGAGCTTCAGTGATCCAGATCATGCCACTGCACTCCAGGCTGG 2166
Db 2112 GTAAACCCGGAGGCGGAGCTTCAGTGATCCAGATCATGCCACTGCACTCCAGGCTGG 2171
Qy 2167 GCGAAGAGCGAGACTCCCGCTCTCAAAAAAATTTTTTGAATAATTCGTAAACCATAC 2226
Db 2172 GCGAAGAGCGAGACTCCCGCTCTCAAAAAAATTTTTTGAATAATTCGTAAACCATAC 2231
Qy 2227 TTTTAAAGATTTTCAAGTGGATTTTAAATAATCTTGTACAGAAATCAGGGTCTTTAGCTA 2286
Db 2232 TTTTAAAGATTTTCAAGTGGATTTTAAATAATCTTGTACAGAAATCAGGGTCTTTAGCTA 2291
Qy 2287 GCAGTCTTCTCCACGCGAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2346
Db 2292 GCAGTCTTCTCCACGCGAGCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 2351
Qy 2347 AATAAATAATATCTCTCTCTGNAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2404
Db 2352 AATAAATAATATCTCTCTCTGNAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAA 2409

RESULT 3
AAA23435

ID AAA23435 standard; cDNA; 2411 BP.

XX AAA23435;

AC AAA23435;

XX 19-JUN-2000 (first entry)

DE cDNA encoding human secreted protein vc38_1, SEQ ID NO:25.

XX Human; secreted protein; cancer; tumour; cardiovascular disorder;

Qy	1141	AAATAGCCAGGAACAAGAAAAACGAGCTGTCAATTATGATGGTGACAGTGGTGCTCTCTTTG	1200
Db	1157	AAATAGCCAGGAACAAGAAAAACGAGCTGTCAATTATGATGGTGACAGTGGTGCTCTCTTTG	1216
Qy	1201	CTGTGTGCTGGGCACCACTTCATGTTGTGCNATATGATGATTGAATCAGTAATATTTGAAA	1260
Db	1217	CTGTGTGCTGGGCACCACTTCATGTTGTGCNATATGATGATTGAATCAGTAATATTTGAAA	1276
Qy	1261	AGGAATATGATGATGTCACAATCAAGATGATTTTTGCTATCGTGCAAAATTATTGGATTTT	1320
Db	1277	AGGAATATGATGATGTCACAATCAAGATGATTTTTGCTATCGTGCAAAATTATTGGATTTT	1336
Qy	1321	CCAACTCCATCTGPAATCCCAATTGTCATGCAATTTATGAATGAAAACTTCAAAAAAAATG	1380
Db	1337	CCAACTCCATCTGPAATCCCAATTGTCATGCAATTTATGAATGAAAACTTCAAAAAAAATG	1396
Qy	1381	TTTTGTCTGCAGTTGTTATTGTCATAGTAATAATAAACCTTCTCTCCAGCACAAAGGCATG	1440
Db	1397	TTTTGTCTGCAGTTGTTATTGTCATAGTAATAATAAACCTTCTCTCCAGCACAAAGGCATG	1456
Qy	1441	GAATTCAGGAATPACAATGATCGGGAAGAAACAAAGTTTTCCTCCAGAGAGAAATCCAG	1500
Db	1457	GAATTCAGGAATPACAATGATCGGGAAGAAACAAAGTTTTCCTCCAGAGAGAAATCCAG	1516
Qy	1501	TGGAGGAAACCAAGGAGAGACATTCAGTGAATGCAACATTTGAAGTCAAAATGTGTGAAC	1560
Db	1517	TGGAGGAAACCAAGGAGAGACATTCAGTGAATGCAACATTTGAAGTCAAAATGTGTGAAC	1576
Qy	1561	AGACAGAGGAGAAAGAAAAAGCTCAAAACACATCTTGCTCTCTTTAGTGCTCGAACTGGCTG	1620
Db	1577	AGACAGAGGAGAAAGAAAAAGCTCAAAACACATCTTGCTCTCTTTAGTGCTCGAACTGGCTG	1636
Qy	1621	AGAAATCTCTTTTAGACAGTGGGCATTAATATTAACCAATATCTTCAATAATTAATGCCCC	1680
Db	1637	AGAAATCTCTCTTTAGACAGTGGGCATTAATATTAACCAATATCTTCAATAATTAATGCCCC	1696
Qy	1681	CAGATTGTAACCCAAAGAGAAAAATTATTTTGACAAAGAGGTCAAAATCTCTTTTAATTCCT	1740
Db	1697	CAGATTGTAACCCAAAGAGAAAAATTATTTTGACAAAGAGGTCAAAATCTCTTTTAATTCCT	1756
Qy	1741	AAGATGATGACAAGAGAAAAACAAATCATGTTTCCATTAAAAAATGACAGAGCTAGTC	1800
Db	1757	AAGATGATGACAAGAGAAAAACAAATCATGTTTCCATTAAAAAATGACAGAGCTAGTC	1816
Qy	1801	CAAGTCAGTGATGTTTCAACCAATNGATCACAATCATTTTAAACANATTTCTGTGTTCCCT	1860
Db	1817	CAAGTCAGTGATGTTTCAACCAATNGATCACAATCATTTTAAACAGATTTCTGTGTTCCCT	1876
Qy	1861	TCTCATTTCCCACTGCTTCACTTGACTAGCCCTTAAAAAGCAACATGAAGGCGCAGGCACG	1920
Db	1877	TCTCATTTCCCACTGCTTCACTTGACTAGCCCTTAAAAAGCAACATGAAGGCGCAGGCACG	1936
Qy	1921	GTGCTCATGCTGTAAATCCAGCACATTTGGGAGGCGCTAGACGGGCGGATCAAGAGTCA	1980
Db	1937	GTGCTCATGCTGTAAATCCAGCACATTTGGGAGGCGCTAGACGGGCGGATCAAGAGTCA	1996
Qy	1981	GGAGATCAAAACCAATCTGGGTAAACGGTGAACACCCATCTCTGCTAAAAAATACAAAAA	2040
Db	1997	GGAGATCAAAACCAATCTGGGTAAACGGTGAACACCCATCTCTGCTAAAAAATACAAAAA	2056
Qy	2041	TTAGCCGGGCGTGTGGCGGCGACCTGTGAGTCCAGCTACTTGGAGGCGCTCAGCGGGAG	2100
Db	2057	TTAGCCGGGCGTGTGGCGGCGACCTGTGAGTCCAGCTACTTGGAGGCGCTCAGCGGGAG	2116
Qy	2101	AATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAATCCGAGATCATGCCACTGCATCCCA	2160
Db	2117	AATGGTGTGAACCCGGGAGGCGGAGCTTGCAGTGAATCCGAGATCATGCCACTGCATCCCA	2176
Qy	2161	GCCTGGGCGAAAGACGAGAGCTCCCCGCTCTCAAAAAAATTTTTTTTGAATAATTCGTAAA	2220
Db	2177	GCCTGGGCGAAAGACGAGAGCTCCCCGCTCTCAAAAAAATTTTTTTTGAATAATTCGTAAA	2236

QY	2221	CCAFACCTTTTAAGATTATTTTCAGTGGATTTTTTAAAAATCTTTGACAGAAATCAGGGTTCT	2228
DB	2237	CCAFACCTTTTAAGATTATTTTCAGTGGATTTTTTAAAAATCTTTGACAGAAATCAGGGTTCT	2296
QY	2281	TAGCTACGAGTTTTTCTCCACACAGTCACCTGTAATGTGACTATGTATTCCTAGATTGAA	2340
DB	2297	TAGCTACGAGTTTTTCTCCACACAGTCACCTGTAATGTGACTATGTATTCCTAGATTGAA	2356
QY	2341	TAGAAAAATAAAATAATATCTTCTCTCTGNAAAAAAATAAAAAAAAAAAAAA	2395
DB	2357	TAGAAAAATAAAATAATATCTTCTCTCTGNAAAAAAATAAAAAAAAAAAAAA	2411
RESULT 4			
ID	ABK14957	standard; DNA; 1710 BP.	
AC	ABK14957;		
DT	08-MAY-2002	(first entry)	
XX	Human neuropeptide Y-like GPCR (NPY-like GPCR) DNA sequence.		
DE			
XX	Human; neuropeptide Y-like G protein-coupled; receptor; antibody;		
KW	NPY-like GPCR; anorectic; antidiabetic; tranquilizer; hypertensive;		
KW	hypotensive; analgesic; cytostatic; antiasthmatic; osteopathic;		
KW	cardiovascular; antiulcer; nootropic; anticonvulsant; antiallergic;		
KW	neuroleptic; antidote; antiparkinsonian; neuroprotective; gene therapy;		
KW	antidepressant; antiasthmatic; diuretic; chromosome 11q12.2; ds.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	misc_feature	394..1689	
FT	/*tag= a		
FT	/note= "Human NPY-like GPCR gene sequence (ABK14956)"		
XX			
PN	WO200204518-A2.		
PP	17-JAN-2002.		
XX			
PF	06-JUL-2001; 2001WO-US21276.		
XX			
PR	06-JUL-2000; 2000US-216523P.		
XX			
PA	(FARB) BAYER CORP.		
XX			
PI	Bloomquist BT, Zhelnin L;		
XX			
DR	WPI; 2002-148166/19.		
XX			
PT	New neuropeptide Y-like G protein-coupled receptor, for detecting		
PT	regulators for preventing, ameliorating, or correcting e.g. obesity,		
PT	anorexia, diabetes, hypotension, hypertension, pain, cancers, or		
PT	neurological disorders -		
XX			
PS	Disclosure; Page 90; 91pp; English.		
XX			
CC	The present invention relates to a new polynucleotide encoding a		
CC	polypeptide neuropeptide Y-like G protein-coupled receptor (NPY-like		
CC	GPCR), comprising a 431 amino acid sequence, given in the specification.		
CC	The polynucleotide of the invention is used to produce a NPY-like GPCR		
CC	polypeptide. The complement of the polynucleotide can be used to detect		
CC	a coding sequence for a NPY-like GPCR. A reagent, preferably an antibody,		
CC	that binds to NPY-like GPCR can be used to detect it. The NPY-like GPCR		
CC	can be used to screen for agents that can regulate the activity of it. A		
CC	reagent that binds to or regulates NPY-like GPCR can be used to reduce		
CC	the biological activity of NPY-like GPCR, or treat a disorder associated		
CC	with abnormal expression of NPY-like GPCR. The reagent is useful for		
CC	preventing, ameliorating or correcting dysfunctions or diseases such as		
CC	obesity, anorexia, bulimia, diabetes, hypotension, hypertension, cocaine		
CC	withdrawal, pain, cancers, asthma, urinary retention, osteoporosis,		
CC	angina pectoris, myocardial infarction, ulcers, allergies, congestive		

CC heart failure, memory enhancement, cardiac and cerebral vasospasm,
 CC pheochromocytoma, ganglioneuroblastoma, Huntington's disease, severe
 CC mental retardation, Alzheimer's disease, Parkinson's disease, psychotic
 CC and neurological disorders including anxiety, schizophrenia, manic
 CC depression, delirium, dementia and dyskinesias. The polypeptide may also
 CC be used in raising specific antibodies. The present nucleic acid sequence
 CC contains the human NYP-like GPCR gene (ABK14956) located on chromosome
 XX 11q12.2.

Sequence 1710 BP; 419 A; 439 C; 429 G; 423 T; 0 other;

Query Match 69.1%; Score 1668.4; DB 24; Length 1710;
 Best Local Similarity 99.9%; Pred. No. 5.1e-255;
 Matches 1669; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 GCCAGAGCGCCAGGACCTAGCGTGGCGCTCCAGACCCAGACCGTGGCGGCCCTCG 60
 Db 41 GCCAGAGCGCCAGGACCTCGCGTGGCGCTCCAGACCCAGACCGTGGCGGCCCTCG 100

61 CTTAGGGAAGAGCAAGGAAGAACTTTATTGGAACCGCAACATTTTGGTCACTGAG 120
 Db 101 CTTAGGGAAGAGCAAGGAAGAACTTTATTGGAACCGCAACATTTTGGTCACTGAG 160

121 ATCGAGTCTCCAGTGCCTTGGCTTCCGCGCTTTATTCGTGGGTTTGAATCCCTGAGCTG 180
 Db 161 ATCGAGTCTCCAGTGCCTTGGCTTCCGCGCTTTATTCGTGGGTTTGAATCCCTGAGCTG 220

181 CTCTCTCTTCCGAACTCCCGGGGTGACGCTAGAGCCCTCCCGCGGGCTGACTCCAG 240
 Db 221 CTCTCTTCCGAACTCCCGGGGTGACGCTAGAGCCCTCCCGCGGGCTGACTCCAG 280

241 AGTAGAAGAGGAGCGGCTCCGCGTGGTCCCGCAAGCCCTCGCTGCCCGCAGATG 300
 Db 281 AGTAGAAGAGGAGCGGCTCCGCGTGGTCCCGCAAGCCCTCGCTGCCCGCAGATG 340

301 CGGATGGCCAGCCAGTACGGCGGTGGCCCGCTCCCGGAGCGCAGCAATGCGAGG 360
 Db 341 CGGATGGCCAGCCAGTACGGCGGTGGCCCGCTCCCGGAGCGCAGCAATGCGAGG 400

361 CGCTTAAACATTAACCCGGAGCAGTTCTCTCGGCTGTGGGAGCACCACCTGAGCGGG 420
 Db 401 CGCTTAAACATTAACCCGGAGCAGTTCTCTCGGCTGTGGGAGCACCACCTGAGCGGG 460

421 AGCAGTTATCGCTCTGTACCGGCTGGCAACCGCTCGTTACACCCAGAGCTGCCGGAC 480
 Db 461 AGCAGTTATCGCTCTGTACCGGCTGGCAACCGCTCGTTACACCCAGAGCTGCCGGAC 520

481 GCGCAAGCTGCCCTCGTGTACCGGCTGGCTATCTTCCGCTGGCGCTTTGGCA 540
 Db 521 GCGCAAGCTGCCCTCGTGTACCGGCTGGCTATCTTCCGCTGGCGCTTTGGCA 580

541 ATGCTCTGGTGTCTACGTGGTGAACCGCAGCAAGCCATGCGCACCTGACCAACATCT 600
 Db 581 ATGCTCTGGTGTCTACGTGGTGAACCGCAGCAAGCCATGCGCACCTGACCAACATCT 640

601 TTATCTGCTCTGGCGCTCAGTGAACCTGCTCATACCTTCTTGTGCATTTCCCGTCA 660
 Db 641 TTATCTGCTCTGGCGCTCAGTGAACCTGCTCATACCTTCTTGTGCATTTCCCGTCA 700

661 TGCTCCAGAACTTCCGCAACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGCAT 720
 Db 701 TGCTCCAGAACTTCCGCAACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGCAT 760

721 TTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATGCTGTGGAAA 780
 Db 761 TTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATGCTGTGGAAA 820

781 GGCACAGGACTGTGCATCTTTTAAATGAAGTGCCAAATACCAACCGAAGGGCTT 840
 Db 821 GGCACAGGACTGTGCATCTTTTAAATGAAGTGCCAAATACCAACCGAAGGGCTT 880

841 TCACAATGCTAGTGTGGTCTGGCTGGTGGCAGTCACTGTAGGATCACCATGTGGCAG 900

Db 881 TCAATGCTAGTGTGTGTCTGTGGTGGCAGTCATCTAGGATCACCCATGTGGCAG 940
 Qy 901 TGCACCACTTGAGATCAATATGACATTCCTATATGAAAAGAAACACATCTGCTGCTTAG 960
 Db 941 TGCACCACTTGAGATCAATATGACATTCCTATATGAAAAGAAACACATCTGCTGCTTAG 1000
 Qy 961 AAGAGTGGACAGCCCTGTGCACAGAGATCTACACACCTTCATCTTGTCTCATCTCT 1020
 Db 1001 AAGAGTGGACAGCCCTGTGCACAGAGATCTACACACCTTCATCTTGTCTCATCTCT 1060

1021 TCCTCTGCTCTTATGGTGAATGCTTATCTGTACAGTAAATTTGGTATGAACCTTTGGA 1080
 Db 1061 TCCTCTGCTCTTATGGTGAATGCTTATCTGTACAGTAAATTTGGTATGAACCTTTGGA 1120

1081 TAAAGAAAAGAGTTGGGATGGTTCAGTCTCGAACTATTTCATGAAAAAGAAATGTCCA 1140
 Db 1121 TAAAGAAAAGAGTTGGGATGGTTCAGTCTCGAACTATTTCATGAAAAAGAAATGTCCA 1180

1141 AAATAGCCAGGAAGAAAACGAGCTGTCAATTTATGATGGTGACAGTGGTCTCTCTTTG 1200
 Db 1181 AAATAGCCAGGAAGAAAACGAGCTGTCAATTTATGATGGTGACAGTGGTCTCTCTTTG 1240

1201 CTGTCTGCTGGGACCAATTCATGCTTGTCCATATGATGATTAATACAGTAATTTTGA 1260
 Db 1241 CTGTCTGCTGGGACCAATTCATGCTTGTCCATATGATGATTAATACAGTAATTTTGA 1300

1261 AGGAATATGATGATGTCACAATCAAGATGATTTTGTCTATCGTCAAAATTTATGGATTT 1320
 Db 1301 AGGAATATGATGATGTCACAATCAAGATGATTTTGTCTATCGTCAAAATTTATGGATTT 1360

1321 CCACCTCATCTGTAATCCCAATCTGTATGCAATTTATGATGAATGAAAACCTTCAAAA 1380
 Db 1361 CCACCTCATCTGTAATCCCAATCTGTATGCAATTTATGATGAATGAAAACCTTCAAAA 1420

1381 TTTTGTCTGCGAGTTGTATTGCTAGTAAATAAACCTTCTCCAGCACAAGGCGATG 1440
 Db 1421 TTTTGTCTGCGAGTTGTATTGCTAGTAAATAAACCTTCTCCAGCACAAGGCGATG 1480

1441 GAAATTCAGGAATTAACATGATGCGGAAGAAAGCAAGTTTCCCTCAGAGAGATCCAG 1500
 Db 1481 GAAATTCAGGAATTAACATGATGCGGAAGAAAGCAAGTTTCCCTCAGAGAGATCCAG 1540

1501 TGGAGGAAACCAAGGAGAGCAATTCAGTGAATGGCAACATTTGAAAGTCAAAATTTGTGA 1560
 Db 1541 TGGAGGAAACCAAGGAGAGCAATTCAGTGAATGGCAACATTTGAAAGTCAAAATTTGTGA 1600

1561 AGACAGAGGAGAGAAAGCTCAACGACATCTTCTCTCTCTTGTAGTCTGAACTGCTG 1620
 Db 1601 AGACAGAGGAGAGAAAGCTCAACGACATCTTCTCTCTCTTGTAGTCTGAACTGCTG 1660

1621 AGAATTCCTCTTTAGACAGTGGGCAATTAATTAACAATATCTTTCATAAT 1670
 Db 1661 AGAATTCCTCTTTAGACAGTGGGCAATTAATTAACAATATCTTTCATAAT 1710

RESULT 5
 AAH49502
 ID AAH49502 standard; DNA; 1296 BP.
 XX
 AC AAH49502;
 XX
 DT 21-SEP-2001 (first entry)
 XX
 DE Human GTP-binding protein-coupled receptor GPRv11 coding sequence.
 XX
 KW GTP-binding protein-coupled receptor; neuroprotective; immunomodulatory;
 KW muscular; urinary; circulatory; anorectic; gene therapy; human;
 XX
 OS guanosine triphosphate; G-protein; ds.
 XX
 PN Homo sapiens.
 XX
 WO200148189-A1.
 XX

XX OS Homo sapiens.
 XX PN WO200078809-A1.
 XX PD 28-DEC-2000.
 XX PF 19-JUN-2000; 2000WO-US16869.
 XX PR 21-JUN-1999; 99US-0337105.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Elshourbagy N, Shabon U;
 XX WPI; 2001-102699/11.
 XX DR P-PDB; AAB48963.
 XX PT New AXOR16, a G-protein coupled receptor polypeptide for screening
 PT agonists and antagonists and for diagnosing and treating microbial
 PT infections, cancer, neurological disorders and asthma -
 XX Claim 2; Page 29; 31pp; English.
 XX The invention relates to the human G protein-coupled receptor AXOR16
 CC (AAB48963) to DNA encoding AXOR16 (AAC87690), and to AXOR16 fragments
 CC and variants. Like all G protein-coupled receptors, AXOR16 has 7
 CC putative transmembrane domains and is involved in signal transduction.
 CC AXOR16 has homology with Gads morhua neuropeptide (NPVBB) F and
 CC homology/ structural similarity with Danio rerio neuropeptide Y (NPYRYA).
 CC The human AXOR16 gene is located on chromosome 11q12.2. The invention
 CC also relates to expression vectors and host cells comprising AXOR16 DNA,
 CC to recombinant expression of AXOR16, and to an AXOR16-specific antibody.
 CC AXOR16 proteins and nucleotides may be used to treat a wide variety of
 CC disorders including bacterial, fungal, protozoal and viral infections,
 CC particularly HIV-1 or HIV-2 infections, pain; cancers; benign prostatic
 CC hypertrophy; diabetes; obesity; anorexia; bulimia; osteoporosis; asthma;
 CC allergies; urinary retention; acute heart failure; hypotension;
 CC hypertension; angina pectoris; myocardial infarction; stroke; ulcers;
 CC migraine; vomiting; psychotic and neurological disorders such as
 CC anxiety, schizophrenia, manic depression, depression, delirium, dementia,
 CC and severe mental retardation, and dyskinesias, such as Parkinson's
 CC disease, Huntington's disease or Gilles de la Tourette's syndrome.
 CC AXOR16 proteins, and nucleotides are useful as vaccines, and AXOR16
 CC proteins, nucleotides and antibodies may be used in screening compounds
 CC for their ability to modulate AXOR16 activity or expression. AXOR16
 CC proteins are also useful for inducing an immunological response in a
 CC mammal against the above diseases, for antibody production, and to
 CC identify membrane bound or soluble receptors for AXOR16. AXOR16
 CC nucleotides are also useful as diagnostic reagents and in chromosome
 CC localisation and tissue expression studies. The present sequence
 CC represents DNA encoding human AXOR16.
 XX SQ Sequence 1296 BP; 344 A; 305 C; 301 G; 346 T; 0 other;

Query Match 53.7%; Score 1296; DB 22; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 3.6e-196;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 ATGACGGCGTTAACATTACCCGGAGACGTTCTCTCGGCTGCTGGGGACCAACCTG 413
 DB 1 ATGACGGCGTTAACATTACCCGGAGACGTTCTCTCGGCTGCTGGGGACCAACCTG 60
 QY 414 ACGGGGAGAGTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 473
 DB 61 ACGGGGAGAGTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 120
 QY 474 CCGGGACGCCCAAGCTGGCCCTCGTGTCTACCGGCTGCTCTACCTTCGCGGCTG 533
 DB 121 CCGGGACGCCCAAGCTGGCCCTCGTGTCTACCGGCTGCTCTACCTTCGCGGCTG 180
 QY 534 TTGGGCAATGCTCTGGTGTCTAGTGGTTCACCGGAGCAAGGCCATGCCACCGTCAAC 593

DB 181 TTGGCAATGCTCTGGTGTCTAGTGGTGTACCGGAGCAAGGCCATGCCACCGTCAAC 240
 QY 594 AACATCTTTATCTGCTCCCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATCCC 653
 DB 241 AACATCTTTATCTGCTCCCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATCCC 300
 QY 654 GTCACCATGCTCTCAGAAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTTTCAAGATG 713
 DB 301 GTCACCATGCTCTCAGAAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTTTCAAGATG 360
 QY 714 GTGCCAATTTGTCTGCTTACCGCTGTGTGTGACAGAAATCTCTCATATGACCTGCATTTGCT 773
 DB 361 GTGCCAATTTGTCTGCTTACCGCTGTGTGTGACAGAAATCTCTCATATGACCTGCATTTGCT 420
 QY 774 GTGGAAGGCACCCAGGGACTTTGTGCATCTTTTAAATGAAGTGCGCAATACACCAACCGA 833
 DB 421 GTGGAAGGCACCCAGGGACTTTGTGCATCTTTTAAATGAAGTGCGCAATACACCAACCGA 480
 QY 834 AGGGCTTTTCACAATGCTAGT 893
 DB 481 AGGGCTTTTCACAATGCTAGT 540
 QY 894 TGGCAGCTGCAACAACTTGAGATCAAAATGATCTTCTATATGAAAAGGAACACATCTGC 953
 DB 541 TGGCAGCTGCAACAACTTGAGATCAAAATGATCTTCTATATGAAAAGGAACACATCTGC 600
 QY 954 TGCTTAGAAGAGTGGACACAGCCCTGTGCACCAAGAGATCTACACCACTTTCATCCTTGTG 1013
 DB 601 TGCTTAGAAGAGTGGACACAGCCCTGTGCACCAAGAGATCTACACCACTTTCATCCTTGTG 660
 QY 1014 ATCTCTTCTCTCTGCTCTTATGT 1073
 DB 661 ATCTCTTCTCTCTGCTCTTATGT 720
 QY 1074 CTTTGCATTAAGAAAGAGTGTGGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1133
 DB 721 CTTTGCATTAAGAAAGAGTGTGGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 QY 1134 ATGTCCAAATAGCCAGGAAGAAACAGAGTGTCTATATGATGTGTGTGTGTGTGTGTGTGTGT 1193
 DB 781 ATGTCCAAATAGCCAGGAAGAAACAGAGTGTCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 840
 QY 1194 CTCTTTGT 1253
 DB 841 CTCTTTGT 900
 QY 1254 TTTTGAAGGAATATGATGATGTCAATCAAGATGATTTTGTCTATCGTGTGCAATTTATTT 1313
 DB 901 TTTTGAAGGAATATGATGATGTCAATCAAGATGATTTTGTCTATCGTGTGCAATTTATTT 960
 QY 1314 GGATTTTCCAACTCCATCTGTATCCCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1373
 DB 961 GGATTTTCCAACTCCATCTGTATCCCATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
 QY 1374 AAAAATGTTTTGT 1433
 DB 1021 AAAAATGTTTTGT 1080
 QY 1434 AGGCATGGAATTCAGGATTTACATGTGCGGAGAAAGCAAGTGTTCCTTCAGAGAG 1493
 DB 1081 AGGCATGGAATTCAGGATTTACATGTGCGGAGAAAGCAAGTGTTCCTTCAGAGAG 1140
 QY 1494 AATCCAGTGGAGAAACCAAAGGAAGCAATTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1553
 DB 1141 AATCCAGTGGAGAAACCAAAGGAAGCAATTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
 QY 1554 TGTGAACAGACAGAGAGAGAAAGGCTCAACAGCATCTTGTCTCTCTTTAGGCTGAA 1613
 DB 1201 TGTGAACAGACAGAGAGAGAAAGGCTCAACAGCATCTTGTCTCTCTTTAGGCTGAA 1260
 QY 1614 CTGGCTGAGAATCTCTCTTTAGACAGTGGGCAATTA 1649
 DB 1261 CTGGCTGAGAATCTCTCTTTAGACAGTGGGCAATTA 1296

RESULT 7	
ABK86287	
ID	ABK86287 standard; cDNA; 1296 BP.
XX	
AC	ABK86287;
DT	27-AUG-2002 (first entry)
XX	
DE	Human TGR346 cDNA.
XX	
KW	Human; TGR346; gene; ss; G-protein coupled receptor; GPCR; TGR;
KW	TGR-associated disorder; signal transduction; renal failure; nephritis;
KW	hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
KW	diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;
KW	sleep disturbance; temperature regulation; blood pressure; hypothalamus;
XX	circadian rhythm.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FH	1..1296
CDS	/tag= a
FT	/product= "Human TGR346 protein"
FT	
XX	
PN	WO200242458-A2.
XX	
PD	30-MAY-2002.
XX	
XX	21-NOV-2001; 2001WO-US43404.
PF	
XX	22-NOV-2000; 2000US-352841P.
PR	12-DEC-2000; 2000US-357636P.
PR	22-JAN-2001; 2001US-361377P.
PR	28-MAR-2001; 2001US-279554P.
PR	29-MAR-2001; 2001US-280696P.
XX	
PA	(TULA-) TULARIK INC.
XX	
PI	Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;
XX	
WPI	; 2002-463633/49.
DR	P-PSDB; AAU97735.
DR	
PT	New isolated G-protein coupled receptor polypeptide, termed TGR, for
PT	diagnosis and treatment of diseases such as renal failure, nephritis,
PT	hypothyroidism, diabetes insipidus, and disturbances of thirst and
PT	sleep -
XX	
PS	Claim 42; Page 74; 98pp; English.
XX	
CC	The invention relates to a G-protein coupled receptor polypeptide (GPCR),
CC	termed TGR, and its associated nucleic acid. The sequences of the
CC	invention are useful for identifying a compound that modulates signal
CC	transduction and for identifying a mammal having a TGR-associated
CC	disorder. The proteins and nucleic acids are useful in diagnosis and
CC	treatment of diseases or conditions such as renal failure, nephritis, and
CC	hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
CC	diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
CC	sleep, temperature regulation, appetite, blood pressure or any other
CC	syndrome or disease associated with the hypothalamus. The sequences can
CC	be used in regulation of circadian rhythms, for use as genetic markers
CC	for the identification of mutations associated with diseases resulting
CC	from GPCR inactivation in particular cell types and for identification of
CC	modulators of GPCR signal transduction. This sequence represents cDNA
CC	encoding the human TGR346 polypeptide.
XX	
SQ	Sequence 1296 BP; 344 A; 305 C; 301 G; 346 T; 0 other;
Query Match	53.7%; Score 1296; DB 24; Length 1296;
Best Local Similarity	100.0%; Pred. No. 3.6e-136;
Matches 1296; Conservative	0; Mismatches 0; Indels 0; Gaps 0;

QY 1434 AGGCATGGAATTCAGGAATTACATGATGCGGAGAAAGCAAAAGTTTCCCTCAGAGAG 1493
 |||||
 Db 1081 AGGCATGGAATTCAGGAATTACATGATGCGGAGAAAGCAAAAGTTTCCCTCAGAGAG 1140
 |||||
 QY 1494 AATCCAGTGGAGGAAACCAAGGAGAACATTCAGTGTATGCAACATGGAAGTCAAAATG 1553
 |||||
 Db 1141 AATCCAGTGGAGGAAACCAAGGAGAACATTCAGTGTATGCAACATGGAAGTCAAAATG 1200
 |||||
 QY 1554 TGTGAACACAGACAGAGGAGAAAGAAAGCTCAAAACGACATCTTGTCTCTCTTTAGTGTCTGAA 1613
 |||||
 Db 1201 TGTGAACACAGACAGAGGAGAAAGAAAGCTCAAAACGACATCTTGTCTCTCTTTAGTGTCTGAA 1260
 |||||
 QY 1614 CTGGCTGAGAAATTCCTTTTACAGAGTGGGATTA 1649
 |||||
 Db 1261 CTGGCTGAGAAATTCCTTTTACAGAGTGGGATTA 1296
 |||||

RESULT 8

ABK14956

ID ABK14956 standard; DNA; 1296 BP.

XX AC

XX AC ABK14956;

XX DT

XX 08-MAY-2002 (first entry)

XX DE

XX Human neuropeptide Y-like GPCR (NPY-like GPCR) gene sequence.

XX KW

KW Human; neuropeptide Y-like G protein-coupled; receptor; antibody; gene;

KW NPY-like GPCR; anorectic; antidiabetic; tranquilizer; hypertensive;

KW hypotensive; analgesic; cytosolic; antispasmodic; osteopathic;

KW cardiovascular; antitumor; neurotropic; anticonvulsant; antiallergic;

KW neuroleptic; antidote; antiparkinsonian; neuroprotective; gene therapy;

KW antidepressant; antispasmodic; diuretic; chromosome 11q12.2; ds.

XX OS

XX Homo sapiens.

XX PH

XX Key Location/Qualifiers

XX CDS

XX 1..1296

XX /*tag= a

XX /product= "Human NPY-like GPCR protein"

XX PN

XX WO200204518-A2.

XX PD

XX 17-JAN-2002.

XX PF

XX 06-JUL-2001; 2001WO-US21276.

XX PR

XX 06-JUL-2000; 2000US-216523P.

XX PA

XX (FARB) BAYER CORP.

XX XX

XX Bloomquist BT, Zhelnin L;

XX PI

XX WPI; 2002-148166/19.

XX DR

XX P-PSDB; ANU76035.

XX XX

XX New neuropeptide Y-like G protein-coupled receptor, for detecting

XX regulators for preventing, ameliorating, or correcting e.g. obesity,

XX anorexia, diabetes, hypotension, hypertension, pain, cancers, or

XX neurological disorders

XX PS

XX Claim 2; Fig 1; 9lpp; English.

XX XX

XX The present invention relates to a new polynucleotide encoding a

XX polypeptide neuropeptide Y-like G protein-coupled receptor (NPY-like

XX GPCR) comprising a 431 amino acid sequence, given in the specification.

XX The polynucleotide of the invention is used to produce a NPY-like GPCR

XX polypeptide. The complement of the polynucleotide can be used to detect

XX a coding sequence for a NPY-like GPCR. A reagent, preferably an antibody,

XX that binds to NPY-like GPCR can be used to detect it. The NPY-like GPCR

XX can be used to screen for agents that can regulate the activity of it. A

XX reagent that binds to or regulates NPY-like GPCR can be used to reduce

XX the biological activity of NPY-like GPCR, or treat a disorder associated

XX CC

XX CC

XX CC

XX CC

CC with abnormal expression of NPY-like GPCR. The reagent is useful for
 CC preventing, ameliorating or correcting dysfunctions or diseases such as
 CC obesity, anorexia, bulimia, diabetes, hypotension, hypertension, cocaine
 CC withdrawal, pain, cancers, asthma, urinary retention, osteoporosis,
 CC angina pectoris, myocardial infarction, ulcers, allergies, congestive
 CC heart failure, memory enhancement, cardiac and cerebral vasospasm,
 CC pheochromocytoma, ganglioneuroblastoma, Huntington's disease, severe
 CC mental retardation, Alzheimer's disease, Parkinson's disease, psychotic
 CC depression, delirium, dementia and dyskinesias. The polypeptide may also
 CC be used in raising specific antibodies. The present nucleic acid sequence
 CC is that of the human NPY-like GPCR gene located on chromosome 11q12.2.
 CC This sequence encodes the human NPY-like GPCR protein of the invention.
 CC This sequence is also included in the human NPY-like GPCR DNA sequence
 CC (ABK14957) of the invention.

XX
 SQ Sequence 1296 BP; 344 A; 305 C; 301 G; 346 T; 0 other;

Query Match 53.7%; Score 1296; DB 24; Length 1296;
 Best Local Similarity 100.0%; Pred. No. 3.6e-196;
 Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 ATGCAGGGCTTAACATTACCCGGAGAGTTCTTCGGCTGCTCGGGACCAACCTG 413

Db 1 ATGCAGGGCTTAACATTACCCGGAGAGTTCTTCGGCTGCTCGGGACCAACCTG 60

QY 414 ACGGGGAGCAGTTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 473

Db 61 ACGGGGAGCAGTTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 120

QY 474 CCGGGACGCGCAAGCTGCGCTGCTCACCGGCTGCTCATCTTCGCTCGGGCTG 533

Db 121 CCGGGACGCGCAAGCTGCGCTGCTCACCGGCTGCTCATCTTCGCTCGGGCTG 180

QY 534 TTGGCAATGCTCTGGGTGTTCTACGTGCTGACCCGCGAGAGGCGCATGCCGCTCACC 593

Db 181 TTGGCAATGCTCTGGGTGTTCTACGTGCTGACCCGCGAGAGGCGCATGCCGCTCACC 240

QY 594 AACATCTTTATCTGCTCTCTTTGGGCTCAGTGACCTGCTCATCTTCTTCTGCAATCCC 653

Db 241 AACATCTTTATCTGCTCTCTTTGGGCTCAGTGACCTGCTCATCTTCTTCTGCAATCCC 300

QY 654 GTACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGTGCTTTCAATTTGCAAGATG 713

Db 301 GTACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGTGCTTTCAATTTGCAAGATG 360

QY 714 GTGCCATTTGTCAGTCTACGGCTGTTGCACAGAAATCTCATATGACCTGCAATGCT 773

Db 361 GTGCCATTTGTCAGTCTACGGCTGTTGCACAGAAATCTCATATGACCTGCAATGCT 420

QY 774 GTGAAAAGGCACCGAGGACTTTGTCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 833

Db 421 GTGAAAAGGCACCGAGGACTTTGTCATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 480

QY 834 AGGCTTTCAATGCTAGGTGGTCTGGCTGGTGGCAGTCACTCGTAGAATCAACCATG 893

Db 481 AGGCTTTCAATGCTAGGTGGTCTGGCTGGTGGCAGTCACTCGTAGAATCAACCATG 540

QY 894 TGGCAGCTGCAACCAACTTGAGATCAATATGCTTCTATATGAAGAGGACACATGCTG 953

Db 541 TGGCAGCTGCAACCAACTTGAGATCAATATGCTTCTATATGAAGAGGACACATGCTG 600

QY 954 TGCTTAGAAGAGTGGACCGACCTGTGTCACAGAAAGATCTACACCACTTCATCTTGTGTC 1013

Db 601 TGCTTAGAAGAGTGGACCGACCTGTGTCACAGAAAGATCTACACCACTTCATCTTGTGTC 660

QY 1014 ATCTCTTCTCTGCTCTTATGGTGAATGCTTATCTGTACAGTAAATTTGGTTATGAA 1073

Db 661 ATCTCTTCTCTGCTCTTATGGTGAATGCTTATCTGTACAGTAAATTTGGTTATGAA 720

QY 1074 CTTTGGATAAAGAAAGAGTTGGGATGTTGAGTGTCTGCAACTATTCTATGGAAGAA 1133

Db 721 CTTTGGATAAAGAAAGAGTTGGGATGTTGAGTGTCTGCAACTATTCTATGGAAGAA 780

[illegible]

XX	Human G protein coupled receptor hRUP4 encoding cDNA SEQ ID NO:39.
DE	XX
XX	XX
XX	Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW	identification; agonist; screening; therapeutic; pharmaceutical;
KW	mutant; ss.
XX	XX
XX	Homo sapiens.
OS	XX
XX	WO2000022131-A2.
PN	XX
XX	XX
XX	20-APR-2000.
PD	XX
XX	13-OCT-1999; 99WO-US24065.
XX	XX
XX	13-OCT-1998; 98US-0170496.
PR	13-NOV-1998; 98US-0108029.
PR	20-NOV-1998; 98US-0109213.
PR	27-NOV-1998; 98US-0110060.
PR	16-FEB-1999; 99US-0120416.
PR	26-FEB-1999; 99US-0121852.
PR	12-MAR-1999; 99US-0123944.
PR	12-MAR-1999; 99US-0123945.
PR	12-MAR-1999; 99US-0123946.
PR	12-MAR-1999; 99US-0123948.
PR	12-MAR-1999; 99US-0123949.
PR	12-MAR-1999; 99US-0123951.
PR	28-MAY-1999; 99US-0136436.
PR	28-MAY-1999; 99US-0136437.
PR	28-MAY-1999; 99US-0136439.
PR	28-MAY-1999; 99US-0137127.
PR	28-MAY-1999; 99US-0137131.
PR	28-MAY-1999; 99US-0137567.
PR	30-JUN-1999; 99US-0141448.
PR	27-AUG-1999; 99US-0151114.
PR	03-SEP-1999; 99US-0152524.
PR	29-SEP-1999; 99US-0156633.
PR	29-SEP-1999; 99US-0156555.
PR	29-SEP-1999; 99US-0156634.
XX	XX
PA	(AREN-) ARENA PHARM INC.
XX	XX
PI	Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
PI	Gore M, Liaw CW, Lin I, Lowitz K, White C;
XX	WPI; 2000-317986/27.
DR	P-PSDB; AAB02843.
DR	XX
XX	XX
PT	Non-endogenous, human G protein-coupled receptors for screening
PT	receptor, inverse or partial agonists useful as therapeutic agents -
XX	Example 1; Page 118-119; 187pp; English.
XX	XX
CC	The present invention describes transmembrane receptors, preferably
CC	human G protein coupled receptors (GPCR), for which the endogenous
CC	ligand is unknown (orphan GPCR receptors). More specifically the present
CC	invention relates to non-endogenous, constitutively activated versions
CC	of a human GPCR. These non-endogenous human GPCRs can be useful for
CC	the direct identification of candidate compounds as receptors agonists,
CC	inverse agonists or partial agonists for use as pharmaceutical agents,
CC	AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
CC	the exemplification of the present invention.
XX	XX
SO	Sequence 1296 BP; 344 A; 304 C; 302 G; 346 T; 0 other:

[illegible]

```
QY 414 AGCGGGAGCATTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 473
Db 61 AGCGGGAGCATTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTG 120
QY 474 CGGGAACGGGCAAGCTGGCCCTCGTGTCTACCGGCTGCTCATCTTCCGCTTGGCGCTC 533
Db 121 CGGGAACGGGCAAGCTGGCCCTCGTGTCTACCGGCTGCTCATCTTCCGCTTGGCGCTC 180
QY 534 TTGTGCAATGCTCTGCTGTCTACGCTGTGACGCTGCGCCGACGAGGCCATGCGACCGTCA 593
Db 181 TTGTGCAATGCTCTGCTGTCTACGCTGTGACGCTGCGCCGACGAGGCCATGCGACCGTCA 240
QY 594 AACATCTTTATCTGCTCTGCGCTGCTGACGCTGCTCATCACCTCTCTCTGCAATCC 653
Db 241 AACATCTTTATCTGCTCTGCGCTGCTGACGCTGCTCATCACCTCTCTCTGCAATCC 300
QY 654 GTCACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGGTCTTTCAATTTGCAAGATG 713
Db 301 GTCACCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGGTCTTTCAATTTGCAAGATG 360
QY 714 GTGCCATTTGTCAGTCTACGCTGCTGTGTGACAGAAATCCTCACTATGACCTGCAATGCT 773
Db 361 GTGCCATTTGTCAGTCTACGCTGCTGTGTGACAGAAATGCTCACTATGACCTGCAATGCT 420
QY 774 GTGGAAGGCAACGAGGACTTTGTGCATCTTTTAAATGAAGTGGCAATACACCAACCGA 833
Db 421 GTGGAAGGCAACGAGGACTTTGTGCATCTTTTAAATGAAGTGGCAATACACCAACCGA 480
QY 834 AGGCTTTTCAATGCTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 893
Db 481 AGGCTTTTCAATGCTAGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 540
QY 894 TGGCAGCTGCAACACTTGCAGATCAATATGACTTCTTATATGAAAGGAAACACATCTGC 953
Db 541 TGGCAGCTGCAACACTTGCAGATCAATATGACTTCTTATATGAAAGGAAACACATCTGC 600
QY 954 TGCTTAGAAGTGGACAGCGCTGTGCAACGAGAGTCTACACCACTTCATCTTGTGTC 1013
Db 601 TGCTTAGAAGTGGACAGCGCTGTGCAACGAGAGTCTACACCACTTCATCTTGTGTC 660
QY 1014 ATCTCTTCTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 1073
Db 661 ATCTCTTCTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 720
QY 1074 CTTTGGATAAAGAAAGTTGGGATGGTTCAAGTCTTCGAACTTATTCATGGAAGAA 1133
Db 721 CTTTGGATAAAGAAAGTTGGGATGGTTCAAGTCTTCGAACTTATTCATGGAAGAA 780
QY 1134 ATGTCCAAATAGCCAGGAAGAAAGAGCTGTCTATATGATGGTGACAGTGGTGGCT 1193
Db 781 ATGTCCAAATAGCCAGGAAGAAAGAGCTGTCTATATGATGGTGACAGTGGTGGCT 840
QY 1194 CTCTTTGCTGTGCTGGGACCACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1253
Db 841 CTCTTTGCTGTGCTGGGACCACTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900
QY 1254 TTTGAAAGGAATATGATGATGTGCATCAAGATGATTTTGTCTATCGTCAAAATATT 1313
Db 901 TTTGAAAGGAATATGATGATGTGCATCAAGATGATTTTGTCTATCGTCAAAATATT 960
QY 1314 GGATTTTCCAACTCACTGTAATCCCAATGCTGATGCTGATGCTGATGCTGATGCTGCT 1373
Db 961 GGATTTTCCAACTCACTGTAATCCCAATGCTGATGCTGATGCTGATGCTGATGCTGCT 1020
QY 1374 AAAAATGTTTGTCTGCAAGTTTGTATTTGCAATGATGATGATGATGATGATGATGAT 1433
Db 1021 AAAAATGTTTGTCTGCAAGTTTGTATTTGCAATGATGATGATGATGATGATGATGAT 1080
QY 1434 AGGCATGGAATTTACGAATTTACATGATGCGGAGGAAGCAAAAGTTTTCCTCAGAGAG 1493
Db 1081 AGGCATGGAATTTACGAATTTACATGATGCGGAGGAAGCAAAAGTTTTCCTCAGAGAG 1140
QY 1494 AATCCAGTGGAGAAACCAAGGAGAGCATTTCAGTGTGATGCAACATGGAAGTCAAA 1553
```

```
Db 1141 AATCCAGTGGAGAAACCAAGGAGAGCATTTCAGTGTGCAACATGGAAGTCAAAATTG 1200
QY 1554 TGTGAACAGACAGAGGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGTCTGAA 1613
Db 1201 TGTGAACAGACAGAGGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGTCTGAA 1260
QY 1614 CTGCTGAGAAATTCCTTTAGACAGTGGGCATTAA 1649
Db 1261 CTGCTGAGAAATTCCTTTAGACAGTGGGCATTAA 1296

RESULT 11
AAF81818
ID AAF81818 standard; cDNA; 1293 BP.
AC AAF81818;
XX
XX 05-JUN-2001 (first entry)
XX Human G protein-coupled receptor protein A027 encoding cDNA SEQ ID NO:2.
XX Human; G protein-coupled receptor; A027; nootropic; neuroprotective;
XX hypotensive; antirheumatic; antiallergic; cardiant; antianigmal;
XX abortifacient; gene therapy; Alzheimer's disease; hypertension;
XX pregnancy termination; rheumatism; allergy; angina pectoris; ss.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1293
XX FT /*tag= a
XX FT /partial
XX FT /product= "A027"
XX FT /note= "G protein coupled receptor; no stop codon given"
XX PN WO200116316-A1.
XX
XX PD 08-MAR-2001.
XX
XX PF 24-AUG-2000; 2000WO-JP05684.
XX
XX PR 27-AUG-1999; 99JP-0241530.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX Watanabe T, Kikuchi K, Shintani Y;
XX
XX WPI; 2001-226689/23.
XX P-PSDB; AAB74773.
XX
XX Human brain-originated guanosine triphosphate protein-coupled receptor,
XX its salt and encoded gene, useful in (gene) diagnosis and development
XX of preventives and remedies for e.g. Alzheimer disease, hypertension
XX and rheumatism
XX
XX Claim 4; Page 78-79; 84pp; Japanese.
XX
XX The present sequence encodes a human G protein-coupled receptor
XX designated A027. A027 has nootropic, neuroprotective, hypotensive,
XX antirheumatic, antiallergic, cardiant, antianigmal and abortifacient
XX activities. A027 can be used as a G protein-coupled receptor protein
XX agonist, G-protein-coupled receptor signal transducer and in gene
XX therapy. A027 can be used in the diagnosis and development of
XX preventives and remedies for diseases associated with dysfunction of
XX A027 e.g. Alzheimer's disease, hypertension, pregnancy termination,
XX rheumatism, allergy and angina pectoris.
XX
XX SQ Sequence 1293 BP; 342 A; 305 C; 301 G; 345 T; 0 other;
Query Match 53.5%; Score 1293; DB 22; Length 1293;
Best Local Similarity 100.0%; Pred. No. 1.1e-195;
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 354 ATGCAGCGCTTAACATTACCCCGAGCAGTTCTCTCGGCTGCTCGGGACCAACAACTTG 413
Db 1 ATGCAGCGCTTAACATTACCCCGAGCAGTTCTCTCGGCTGCTCGGGACCAACAACTTG 60
Qy 414 ACGGGGAGAGTTTCATGCTCTGTACCGGCTGGGACCGCTCGTGTACACCCAGAGTG 473
Db 61 ACGGGGAGAGTTTCATGCTCTGTACCGGCTGGGACCGCTCGTGTACACCCAGAGTG 120
Qy 474 CCGGAGCGCGCAAGCTTGGCCCTCTGTGCTCAACCGGCTGTCTATCTTCGCCCTCGCGCTC 533
Db 121 CCGGAGCGCGCAAGCTTGGCCCTCTGTGCTCAACCGGCTGTCTATCTTCGCCCTCGCGCTC 180
Qy 534 TTTGGCAATGCTCTGGTGTCTACCGCTGTGACCGGCAAGGCAATGCAAGATG 593
Db 181 TTTGGCAATGCTCTGGTGTCTACCGCTGTGACCGGCAAGGCAATGCAAGATG 240
Qy 594 AACATCTTTATCTGCTCTTGGCGTCACTGACCTGCTCATCACTTCTTGTGATTCCTCC 653
Db 241 AACATCTTTATCTGCTCTTGGCGTCACTGACCTGCTCATCACTTCTTGTGATTCCTCC 300
Qy 654 GTACCAATGCTTCCAGACATTCGACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 713
Db 301 GTACCAATGCTTCCAGACATTCGACAACTGGCTGGGGGTGCTTTTCATTTGCAAGATG 360
Qy 714 GTGCCATTTGTCCAGTCTACCGCTGTGTGACAGAAATCTCACTATGACCTGCAATGCT 773
Db 361 GTGCCATTTGTCCAGTCTACCGCTGTGTGACAGAAATCTCACTATGACCTGCAATGCT 420
Qy 774 GTGAAAGGACACGAGGACCTGTGCACTCTTTTAAATGAAGTGGCAATACCAACCGA 833
Db 421 GTGAAAGGACACGAGGACCTGTGCACTCTTTTAAATGAAGTGGCAATACCAACCGA 480
Qy 834 AGGCTTTTCAATGCTAGTGTGTCTGGCTGGGAGTCATCGTAGGATCAACCATG 893
Db 481 AGGCTTTTCAATGCTAGTGTGTCTGGCTGGGAGTCATCGTAGGATCAACCATG 540
Qy 894 TGGCACGTGCAACAACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGC 953
Db 541 TGGCACGTGCAACAACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGC 600
Qy 954 TGCTTAGAGTGGACAGCCCTGTGACAGAGATCTACACCACTTCACTCTGTC 1013
Db 601 TGCTTAGAGTGGACAGCCCTGTGACAGAGATCTACACCACTTCACTCTGTC 660
Qy 1014 ATCTCTTCTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATGTTATGAA 1073
Db 661 ATCTCTTCTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATGTTATGAA 720
Qy 1074 CTTTGGATAAGAAAGAGTTGGGGATGGTTCACTGCTTCACTTCACTGAAAGAA 1133
Db 721 CTTTGGATAAGAAAGAGTTGGGGATGGTTCACTGCTTCACTTCACTGAAAGAA 780
Qy 1134 ATGTCCAAATAGCAGAGAGAGAGAGAGCTGTCTATGATGATGAGTGGTGGCT 1193
Db 781 ATGTCCAAATAGCAGAGAGAGAGAGAGCTGTCTATGATGATGAGTGGTGGCT 840
Qy 1194 CTCTTTGCTGTGCTGGGACCACTTCCATGCTTCCATATGATGATGATGATGAT 1253
Db 841 CTCTTTGCTGTGCTGGGACCACTTCCATGCTTCCATATGATGATGATGATGAT 900
Qy 1254 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
Db 901 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
Qy 1314 GGATTTTCCAACTCCATCTGTAATCCCATGCTATGCAATTTGATGATGATGATGAT 1373
Db 961 GGATTTTCCAACTCCATCTGTAATCCCATGCTATGCAATTTGATGATGATGATGAT 1020
Qy 1374 AAAATGTTTGTGCTGCAAGTTTGTATGATGATGATGATGATGATGATGATGAT 1433
Db 1021 AAAATGTTTGTGCTGCAAGTTTGTATGATGATGATGATGATGATGATGATGAT 1080

Qy 1434 AGGCATGGAATTCAGGAATTACAATGATCGGAAGAAAGCAAGTTTTCCTCAGAGAG 1493
Db 1081 AGGCATGGAATTCAGGAATTACAATGATCGGAAGAAAGCAAGTTTTCCTCAGAGAG 1140
Qy 1494 AATCCAGTGGAGAAACCAAGGAGAGCAATTCAGTGTGCAACATGGAAGTCAAAATG 1553
Db 1141 AATCCAGTGGAGAAACCAAGGAGAGCAATTCAGTGTGCAACATGGAAGTCAAAATG 1200
Qy 1554 TGTGAACAGACAGAGAGAGAGAGAGAGAGTCAAAAGACATCTTGTCTCTTTAGTCTGAA 1613
Db 1201 TGTGAACAGACAGAGAGAGAGAGAGAGTCAAAAGACATCTTGTCTCTTTAGTCTGAA 1260
Qy 1614 CTGGCTGAGAATTCCTCTTTAGACAGTGGGCAT 1646
Db 1261 CTGGCTGAGAATTCCTCTTTAGACAGTGGGCAT 1293
RESULT 12
AAA46115
ID AAA46115 standard; cDNA; 1296 BP.
XX AC AAA46115;
XX AC
XX AC
DT 22-AUG-2000 (first entry)
XX Human G protein coupled receptor hRUP4 (V272K) cDNA SEQ ID NO:127.
DE Human; G protein coupled receptor; GPCR; transmembrane receptor;
KW identification; agonist; screening; therapeutic; pharmaceutical;
KW mutant; ss.
XX Homo sapiens.
OS Synthetic.
XX WO200022131-A2.
DN 20-APR-2000.
PD 13-OCT-1999; 99WO-US24065.
PP 13-OCT-1998; 98US-0170496.
PP 12-NOV-1998; 98US-0108029.
PP 20-NOV-1998; 98US-0109213.
PP 27-NOV-1998; 98US-0110060.
PR 16-FEB-1999; 99US-0120416.
PR 26-FEB-1999; 99US-0121852.
PR 12-MAR-1999; 99US-0123944.
PR 12-MAR-1999; 99US-0123945.
PR 12-MAR-1999; 99US-0123946.
PR 12-MAR-1999; 99US-0123948.
PR 12-MAR-1999; 99US-0123949.
PR 12-MAR-1999; 99US-0123951.
PR 28-MAY-1999; 99US-0136436.
PR 28-MAY-1999; 99US-0136437.
PR 28-MAY-1999; 99US-0136439.
PR 28-MAY-1999; 99US-0137127.
PR 28-MAY-1999; 99US-0137131.
PR 30-JUN-1999; 99US-0137567.
PR 27-AUG-1999; 99US-0151114.
PR 03-SEP-1999; 99US-0152524.
PR 29-SEP-1999; 99US-0156633.
PR 29-SEP-1999; 99US-0156555.
PR 29-SEP-1999; 99US-0156634.
XX (AREN-) ARENA PHARM INC.
PA Behan DP, Lehmann-Bruinsma K, Chalmers DT, Chen R, Dang HT;
XX Gore M, Liaw CW, Lin I, Lowitz K, White C;
PI WPI; 2000-317986/27.
XX P-PSDB; AAB02853.
XX

PT Non-endogenous, human G protein-coupled receptors for screening
 PT receptor, inverse or partial agonists useful as therapeutic agents -
 XX Example 2; Page 163-164; 187pp; English.

CC The present invention describes transmembrane receptors, preferably
 CC human G protein coupled receptors (GPCR), for which the endogenous
 CC ligand is unknown (orphan GPCR receptors). More specifically the present
 CC invention relates to non-endogenous, constitutively activated versions
 CC of a human GPCR. These non-endogenous human GPCRs can be useful for
 CC the direct identification of candidate compounds as receptors agonists,
 CC inverse agonists or partial agonists for use as pharmaceutical agents.
 CC AAA46017 to AAA46126 and AAB02825 to AAB02859 represent sequences used in
 CC the exemplification of the present invention.

XX Sequence 1296 BP; 346 A; 303 C; 302 G; 345 T; 0 other;

Query Match 53.4%; Score 1289.6; DB 21; Length 1296;
 Best Local Similarity 99.7%; Pred. No. 3.7e-195;
 Matches 1292; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 354 ATGCAGGCGCTTAAACATTACCCCGGAGCAGTCTCTCGGCTGCTCGGACCAACCTG 413
 Db 1 ATGCAGGCGCTTAAACATTACCCCGGAGCAGTCTCTCGGCTGCTCGGACCAACCTG 60

QY 414 ACGGGGAGCAGTTTCATCGCTCTGTACCGGCTGCGACCGCTCTGTACACCCGAGCTG 473
 Db 61 ACGGGGAGCAGTTTCATCGCTCTGTACCGGCTGCGACCGCTCTGTACACCCGAGCTG 120

QY 474 CCGGACCGGCAAGCTGGCCCTCTGTCTACCGGCTGCTCACTTGGCCCTGGCGTCC 533
 Db 121 CCGGACCGGCAAGCTGGCCCTCTGTCTACCGGCTGCTCACTTGGCCCTGGCGTCC 180

QY 534 TTGGCAATGCTCTGCTGTCTTACCTGTGTGACCGCAGCAGGCGCATCGCAGCTCACC 593
 Db 181 TTGGCAATGCTCTGCTGTCTTACCTGTGTGACCGCAGCAGGCGCATCGCAGCTCACC 240

QY 594 AACATCTTTATCTGCTCTCTGGCGCTCAGTCACTGCTCATCACCTTCTTCTGCAATCCC 653
 Db 241 AACATCTTTATCTGCTCTCTGGCGCTCAGTCACTGCTCATCACCTTCTTCTGCAATCCC 300

QY 654 GTCCACCATGCTCCAGAACATTCGACAACTGGCTGGGGGGTGTCTTCAATTTGCAAGATG 713
 Db 301 GTCCACCATGCTCCAGAACATTCGACAACTGGCTGGGGGGTGTCTTCAATTTGCAAGATG 360

QY 714 GTGCCAATTTGCTCAAGTCTTACCGCTCTTGTGACAGAAATCCCTCACTATCACTGCAATGCT 773
 Db 361 GTGCCAATTTGCTCAAGTCTTACCGCTCTTGTGACAGAAATGCTCACTATCACTGCAATGCT 420

QY 774 GTGAAAGGCAACAGGACTTGTGCATCTTTTAAATGAAATGAGTGGCAATACCAACCGA 833
 Db 421 GTGAAAGGCAACAGGACTTGTGCATCTTTTAAATGAAATGAGTGGCAATACCAACCGA 480

QY 834 AGGCTTTTCAAGTGTAGTGTGTCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 893
 Db 481 AGGCTTTTCAAGTGTAGTGTGTCTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 540

QY 894 TGGCAGCTGCAACACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGCG 953
 Db 541 TGGCAGCTGCAACACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGCG 600

QY 954 TGCTTAGAAGAGTGGAACAGCCCTGTGCAACAGAGATCTACACCACCTTCACTCTTGTGTC 1013
 Db 601 TGCTTAGAAGAGTGGAACAGCCCTGTGCAACAGAGATCTACACCACCTTCACTCTTGTGTC 660

QY 1014 ATCCCTTCT 1073
 Db 661 ATCCCTTCT 720

QY 1074 CTTTGGATAAAGAAAGAGTGGGATGTTTGTGAGTGTGCACTTATTCATGGAAGAA 1133
 Db 721 CTTTGGATAAAGAAAGAGTGGGATGTTTGTGAGTGTGCACTTATTCATGGAAGAA 780

QY 1134 ATGTCCAAATAGCCAGGAGAAAGACGAGCTGCTATTATGATGGTGCAGTGGTGGCT 1193
 Db 781 ATGTCCAAATAGCCAGGAGAAAGACGAGCTGCTATTATGATGGTGCAGTGGTGGCT 840

QY 1194 CTCCTTGTGTGTCTGGGACCACTTCCATGTTCTCCATATGATGATTTGAATACAGTAAT 1253
 Db 841 CTCCTTGTGTGTCTGGGACCACTTCCATGTTCTCCATATGATGATTTGAATACAGTAAT 900

QY 1254 TTTGAAAGGAATATGATGATGTCACAAATCAAGATGATTTTGTCTATCGTCAAAATATT 1313
 Db 901 TTTGAAAGGAATATGATGATGTCACAAATCAAGATGATTTTGTCTATCGTCAAAATATT 960

QY 1314 GGATTTTCCCACTCCATCTGTGATCCCATCTGCTATGATGATTTATGAAATGAAACTTCAA 1373
 Db 961 GGATTTTCCCACTCCATCTGTGATCCCATCTGCTATGATGATTTATGAAATGAAACTTCAA 1020

QY 1374 AAAAATGTTTGTCTGCACTTGTATTGTCATAGTAATAAATAAACCTTCTCTCCAGCAAA 1433
 Db 1021 AAAAATGTTTGTCTGCACTTGTATTGTCATAGTAATAAATAAACCTTCTCTCCAGCAAA 1080

QY 1434 AGGCATGGAATTCAGGAATTAATGATGCGGAGAAAGCAAAAGTTTTCCCTCAGAGAG 1493
 Db 1081 AGGCATGGAATTCAGGAATTAATGATGCGGAGAAAGCAAAAGTTTTCCCTCAGAGAG 1140

QY 1494 AATCCAGTGGAGAAACCAAGGAGAGAGCATTCAGTGTGCAACATTTGAAGTCAAAATTG 1553
 Db 1141 AATCCAGTGGAGAAACCAAGGAGAGAGCATTCAGTGTGCAACATTTGAAGTCAAAATTG 1200

QY 1554 TGTGAAACAGACAGAGGAGAAAGAAAGCTCAACAGCATCTTGTCTCTCTTTAGTCTGAA 1613
 Db 1201 TGTGAAACAGACAGAGGAGAAAGAAAGCTCAACAGCATCTTGTCTCTCTTTAGTCTGAA 1260

QY 1614 CTGGCTCAGAAATCTCTCTTTAGACAGTGGGCAATTA 1649
 Db 1261 CTGGCTCAGAAATCTCTCTTTAGACAGTGGGCAATTA 1296

RESULT 13
 ABZ42861
 ID ABZ42861 standard; DNA; 1368 BP.
 XX AC ABZ42861;
 XX AC
 DT 04-MAR-2003 (first entry)
 XX
 DE Human G protein-coupled receptor GPR103 nucleotide SEQ ID NO:509.
 XX
 KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;
 KW G protein-coupled receptor modulator; antibody; immune-related disease;
 KW growth-related disease; cell regeneration-related disease; AIDS; cancer;
 KW immunological-related cell proliferative disease; autoimmune disease;
 KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;
 KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;
 KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;
 KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;
 KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;
 KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;
 KW ulcer; gene; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO200261087-A2.
 XX
 PD 08-AUG-2002.
 XX
 PF 19-DEC-2001; 2001WO-US50107.
 XX
 PR 19-DEC-2000; 2000US-257144P.
 XX
 PA (LIFE-) LIFESPAN BIOSCIENCES INC.
 XX
 PI Burmer GC, Roush CL, Brown JP;
 XX

DR WPI; 2003-046718/04.
DR P-PSDB; ABP82011.

PT New isolated antigenic peptides e.g., for G protein-coupled receptors
PT (GPCR), useful for diagnosing and designing drugs for treating
PT conditions in which GPCRs are involved, e.g. AIDS, Alzheimer's disease,
PT cancer or autoimmune diseases

XX Disclosure; Fig 1; 523pp; English.

XX The present invention describes antigenic peptides (I) comprising:
CC (a) any one of 1601 sequences (see ABP82019 to ABP83619) of 12-24 amino
CC acids, also described: (1) an assay for the detection of a particular
CC G protein-coupled receptor (GPCR) or a candidate polypeptide in a sample;
CC and (2) an isolated antibody having high specificity and high affinity
CC or avidity for a particular GPCR. (I) can be used as GPCR modulators and
CC in gene therapy. The antigenic peptides for GPCRs are useful in detecting
CC an antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension,
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. AB242523 to AB242869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention.

SQ Sequence 1368 BP; 345 A; 346 C; 319 G; 358 T; 0 other;

Query Match 44.7%; Score 1078.4; DB 25; Length 1368;
Best Local Similarity 91.3%; Pred. No. 8.9e-162;
Matches 1213; Conservative 0; Mismatches 1; Indels 114; Gaps 2;

QY 322 GCGGTGGCCCGCTCCCGGAGCGGACACAGCATGAGCGGCTTAACATTACCCGGAGC 381
DB 155 GCGGTGGCCCGCTCCCGGAGCGGACACAGCATGAGCGGCTTAACATTACCCGGAGC 214
QY 382 AGTTCTCTCGGTGCTGCGGACCAACCTGACGCGGAGGAGTTCATCGCTCTGACC 441
DB 215 AGTTCTCTCGGTGCTGCGGACCAACCTGACGCGGAGGAGTTCATCGCTCTGACC 274
QY 442 GCGTGGACCGGCTGCTTACACCCAGAGTGGCGGACGCGCAAGCTGGCCCTCGTGC 501
DB 275 GCGTGGACCGGCTGCTTACACCCAGAGTGGCGGACGCGCAAGCTGGCCCTCGTGC 334
QY 502 TCACCGGCGTGTCTATCTTCGCGCTGGCGCTCTTGGCAATGCTCTGGTCTACGTTG 561
DB 335 TCACCGGCGTGTCTATCTTCGCGCTGGCGCTCTTGGCAATGCTCTGGTCTACGTTG 394
QY 562 TCACCGGACGAGGCGCATGCGCACCGTCCACCAACATCTTATCTGCTCTCTGGGCTCA 621
DB 395 TCACCGGACGAGGCGCATGCGCACCGTCCACCAACATCTTATCTGCTCTCTGGGCTCA 454
QY 622 GTGACTGCTCATCAGTCTCTTCTGATTCGCGTACCAATGCTCCAGAACATTTCCGACA 681
DB 455 GTGACTGCTCATCAGTCTCTTCTGATTCGCGTACCAATGCTCCAGAACATTTCCGACA 514
QY 682 ACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGGCAATTTGTCAGTCTACGCTGTTG 741
DB 515 ACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGGCAATTTGTCAGTCTACGCTGTTG 574
QY 742 TGACAGAAATCTCTCACTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 801
DB 575 TGACAGAAATCTCTCACTATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 634

QY 802 CTTTAAATGAAGTGGCAATACACCAACCGAAGGGCTTTCAATGCTAGGTGGTCT 861
DB 635 CTTTAAATGAAGTGGCAATACACCAACCGAAGGGCTTTCAATGCTAGGTGGTCT 694
QY 862 GCGTGGTGGCAGTCATCGTAGGATCACCCATGTGGCAGCGTCAACAACTTGAGATCAAT 921
DB 695 GCGTGGTGGCAGTCATCGTAGGATCACCCATGTGGCAGCGTCAACAACTTGAGATCAAT 754
QY 922 ATGACTTCTTATATGAAGAAGAACACATCTGCTGCTTAGAAGTGGACAGCCCTGTGC 981
DB 755 ATGACTTCTTATATGAAGAAGAACACATCTGCTGCTTAGAAGTGGACAGCCCTGTGC 814
QY 982 ACCAGAGATCTACACCACTTCACTCTTGTCTCATCTCTCTCTCTCTCTCTCTCTCT 1041
DB 815 ACCAGAGATCTACACCACTTCACTCTTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 868
QY 1042 TGCTTATCTGTACAGTAAATTTGTTATGAATTTGGATAAAGAAAGAGTTGGGATG 1101
DB 869 ----- 868
QY 1102 GTTCAGTGTTCGAACATTAATTCAGTAAAGAAATGTCAAAATAGCCAGAAAGAAAC 1161
DB 869 -----GGAAGAAAGAAAC 880
QY 1162 GAGCTGTCAATTAATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1221
DB 881 GAGCTGTCAATTAATGATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 940
QY 1222 AGTTGTCTCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1281
DB 941 ATGTTGTCTCATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 1000
QY 1282 TCAAGATGATTTTGTCTATCGTGGCAATTAATTTGGATTTTCCAACTCCATCTGTAATCCCA 1341
DB 1001 TCAAGATGATTTTGTCTATCGTGGCAATTAATTTGGATTTTCCAACTCCATCTGTAATCCCA 1060
QY 1342 TTGCTATGCTATTAATGAATGAAACTTTCAAAAAAATGTTTTGCTGTCAGTTGTTATT 1401
DB 1061 TTGCTATGCTATTAATGAATGAAACTTTCAAAAAAATGTTTTGCTGTCAGTTGTTATT 1120
QY 1402 GCAATGTAATAAATACCTTCTCTCCAGCACAAAGGCATGGAAATTCAGGAATTAACAATCA 1461
DB 1121 GCATAGTAATAAATACCTTCTCTCCAGCACAAAGGCATGGAAATTCAGGAATTAACAATCA 1180
QY 1462 TCGCGAAGAAAGCAAGTCTTCCCTCAGAGAGAAATCCAGTGGAGAAACCAAGAGAGAG 1521
DB 1181 TCGCGAAGAAAGCAAGTCTTCCCTCAGAGAGAAATCCAGTGGAGAAACCAAGAGAGAG 1240
QY 1522 CATTGAGTGGTGGCAACATTAAGTCAAAATTTGTGAACAGACAGAGAGAGAAAGAGC 1581
DB 1241 CATTGAGTGGTGGCAACATTAAGTCAAAATTTGTGAACAGACAGAGAGAGAAAGAGC 1300
QY 1582 TCACACGACATCTGCTCTCTTTAGTCTGAATCGGCTGAGAAATCTCTTTAGACAGTG 1641
DB 1301 TCACACGACATCTGCTCTCTTTAGTCTGAATCGGCTGAGAAATCTCTTTAGACAGTG 1360
QY 1642 GGCATTAA 1649
DB 1361 GGCATTAA 1368

RESULT 14

ABK86292

ID ABK86292 standard; cDNA; 1737 BP.

XX ABK86292;

AC ABK86292;

XX 27-AUG-2002 (first entry)

XX Mouse TGR346a cDNA.

DE Mouse; TGR346a; gene; ss; G-protein coupled receptor; GPCR; TGR;

XX Mouse; TGR346a; gene; ss; G-protein coupled receptor; GPCR; TGR;

TGR-associated disorder; signal transduction; renal failure; nephritis; hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder; diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite; sleep disturbance; temperature regulation; blood pressure; hypothalamus; circadian rhythm.

OS Mus sp.

Key Location/Qualifiers
 CDS 1..1302
 FT /*tag= a
 FT /product= "Mouse TGR346a protein"

XX WO200242458-A2.

XX 30-MAY-2002.

XX 21-NOV-2001; 2001WO-US43404.

XX 22-NOV-2000; 2000US-252841P.

XX 22-DEC-2000; 2000US-257636P.

XX 12-JAN-2001; 2001US-261377P.

XX 28-MAR-2001; 2001US-279554P.

XX 29-MAR-2001; 2001US-280696P.

XX (TULA-) TULARIK INC.

XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai K, Gupte JS;

XX WPI; 2002-463633/49.

XX P-PSDB; AAU97740.

XX New isolated G-protein coupled receptor polypeptide, termed TGR, for

XX diagnosis and treatment of diseases such as renal failure, nephritis,

XX hypothyroidism, diabetes insipidus, and disturbances of thirst and

XX sleep.

XX Claim 22; Page 77-78; 98pp; English.

XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),
 CC termed TGR, and its associated nucleic acid. The sequences of the
 CC invention are useful for identifying a compound that modulates signal
 CC transduction and for identifying a mammal having a TGR-associated
 CC disorder. The proteins and nucleic acids are useful in diagnosis and
 CC treatment of diseases or conditions such as renal failure, nephritis,
 CC hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,
 CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
 CC sleep, temperature regulation, appetite, blood pressure or any other
 CC syndrome or disease associated with the hypothalamus. The sequences can
 CC be used in regulation of circadian rhythms, for use as genetic markers
 CC for the identification of mutations associated with diseases resulting
 CC from GPCR inactivation in particular cell types and for identification of
 CC modulators of GPCR signal transduction. This sequence represents cDNA
 CC encoding the mouse TGR346a polypeptide.

XX Sequence 1737 BP; 473 A; 384 C; 392 G; 488 T; 0 other;

Query Match 41.2%; Score 994.6; DB 24; Length 1737;
 Best Local Similarity 81.2%; Pred. No. 1.6e-148;
 Matches 1172; Conservative 0; Mismatches 259; Indels 12; Gaps 1;

QY 354 ATGCGGCGCTTACATTACCCGGAGCAGTTCTCTCGGCTGCTGGGGACCAACACCTG 413
 DB 1 ATGCGGCGCTTACATCAACCGGAGCAGTTTCCCGGCTGCTGGAGCGGCAACCTG 60

QY 414 ACGGGGAGCAGTTTCATCGCTCTGTACCGCTGCGACCGCTGCTACACCCCGAGAGCTG 473
 DB 61 ACTCGGGAACAGTTCATTATCGCTATGCGCTGCGACCGCTGCTACATCCGGAGCTG 120

QY 474 CCGGGAGCGCCCAAGCTGGCCCTCGTCTCACCGGCGTGTCTATCTTGGCCCTGGGCGCTC 533
 DB 121 CCGGCGCGCTTAACTGGGCTTTGGCTGGCTGGAGCACTCATTTTGGCCCTGGGCGCTC 180

QY 534 TTGGCAATGCTCTGGTGTTCTTACGTGGTGACCGGAGCAAGCCATGCGCAGCGTCACC 593
 DB 181 TTGGCAATGCTCTGGTGTTCTTACGTGGTGACCGGAGCAAGCCATGCGCAGCGTCACC 240
 QY 594 AACATCTTTATCTGCTCTTTGGGCGCTCAGTGACCTGCTCATACCTTCTTCTGCAATCCC 653
 DB 241 AACATCTTCTGCTCTCTGCGCACTCAGTGATCTGCTCATTCGCTTCTTCTGCAATCCC 300
 QY 654 GTCACCATGCTCCAGAACATTTCCGCAACTGCTGGGGGCTGCTTCAATGGAAGATG 713
 DB 301 GTCACCATGCTCCAGAACATTTCCGCAACTGCTGGGGGCTGCTTCAATGGAAGATG 360
 QY 714 GTGCCATTTGTCACAGTCTACCGCTGTTGTACAGAAATCTCACTATGACTGACTGCT 773
 DB 361 GTGCCCTTCGTCAGTCCACTGCTGTTGTGACGGAATCTCTACCATGACTTGCATGCT 420
 QY 774 GTGGAAGGCAACAGGAGCTTGTGCATTCCTTTTAAATGAAGTGGCAATACACCAACCGA 833
 DB 421 GTTGAGAGGCAACAGGAGCTTGTGCATTCCTTTTAAATGAAGTGGCAATACACCAACCGA 480
 QY 834 AGGCTTTTACCAATGCTAGTGTGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 893
 DB 481 AGGCTTTTACCAATGCTAGTGTGCTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 540
 QY 894 TGGCAGTGCACAACTTGAGATCAAATATGACTTCTCTATATGAAAAGGAACACATCTGC 953
 DB 541 TGGCAGTGCACAACTTGAGATCAAATATGACTTCTCTATATGAAAAGGAACACATCTGC 600
 QY 954 TGGTTAGAGAGTGGACACGCGCTGTGACCAAGAGATCTACACACCTTCACTCTGTC 1013
 DB 601 TGGTTAGAGAGTGGACACGCGCTGTGACCAAGAGATCTACACACCTTCACTCTGTC 660
 QY 1014 ATCTCTTCTCTCTGCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATGCTGTATGAA 1073
 DB 661 ATCTCTTCTCTCTGCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATGCTGTATGAA 720
 QY 1074 CTGTGATAAAGAAAGAGTGGGATGGTTTCACTGCTTTCGAACTATTTCATGAAAAGAA 1133
 DB 721 CTGTGATAAAGAAAGAGTGGGATGGTTTCACTGCTTTCGAACTATTTCATGAAAAGAA 780
 QY 1134 ATGTCCAAATAGCCAGGAGAGAAACGAGCTGCTATGATGTGTGACAGTGGTGGCT 1193
 DB 781 ATGTCCAAATAGCCAGGAGAGAAACGAGCTGCTATGATGTGTGACAGTGGTGGCT 840
 QY 1194 CTCTTGTGCTGTGCTGGGCGCCATTCCTGCTGCTCCATGCTCCATGATGATGATGATGAT 1253
 DB 841 CTCTTGTGCTGTGCTGGGCGCCATTCCTGCTGCTCCATGCTCCATGATGATGATGATGAT 900
 QY 1254 TTGAAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
 DB 901 TTGAAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 1314 GGATTTTCCAACTCCATCTGTAATCCCATCTGCTATGCTATGATGATGATGATGATGATGAT 1373
 DB 961 GGCTTTTCCAACTCCATCTGTAATCCCATCTGCTATGCTATGATGATGATGATGATGATGAT 1020
 QY 1374 AAAATGTTTTTGTCTGCACTTGTGTTATTCATAGTAAATATAAACCTTCTCTCCAGACAA 1433
 DB 1021 AAGAATTTTTTGTCTGCGGTTGTTATTCATAGTAAAGAGAAACCTTCTCTCCAGACAA 1080
 QY 1434 AGGCATGGAATATCAGGAATTAATGATGCGGAAGAAAGCAAGTTCCTTCAGAGAG 1493
 DB 1081 AAGCCTGGAAATTCCTGGGATTTCAATGATGCAAAAGAGAGCAAGTTCATCAGATCACAG 1140
 QY 1494 AATCCAGTGGAGAAACCAAGGAGAGCATTCAGTGTGGCAACATTCAGATGCAAAATG 1553
 DB 1141 CGTCAGTGGGGAAGCCAAAGGAGCATTAATTCAGCGATGCGCAACGTTGATGCAAAATG 1200
 QY 1554 TGTGAACACAGAGAGAGAGAAAGAAAGCTCAACACGACATCTTGTCTCTCTTTAGGCTGAA 1613
 DB 1201 TGTGAGCAGCAGGAGAGAGAAAGGCACTCAAGCCAGAGCTTGCCTCTCTTTAGTCTGAA 1260
 QY 1614 CTGGCTGAGAAATTCCTCTTTAGACAGTGGGCAATTAATTATACAAATATCTTCAATAA 1673

Db 1261 CTTTCTGAAACTTACTTTTCGGCAGTGGACATGAATGTATGATATCCTCATAGCTAA 1320
Qy 1674 TGCCCTTCAGATTGTAACCCAAAGAGAAAATTTTTCGAGCAAGGTCAAAATCTCTTTT 1733
Db 1321 TATCATTTGTAT-----GGAAAGTATTTTAAGCAAGGTGAGCACTATTTT 1368
Qy 1734 AATTCCTTAAGTATGACAAAGAGAAAACAAATCATGTTTCCATTTAAATAATGACAGAG 1793
Db 1369 TTTAAATGACAAAGAGAAAACAGACATGTTTCCATTTAAATGAACATAATACATAAC 1428
Qy 1794 GCT 1796
Db 1429 ACT 1431
RESULT 15
ABK86293
ID ABK86293 standard; cDNA; 1772 BP.
XX
AC ABK86293;
AT
DT 27-AUG-2002 (first entry)
XX
DE Mouse TGR346b cDNA.
XX
KW Mouse; TGR346b; gene; ss; G-protein coupled receptor; GPCR; TGR;
KW TGR-associated disorder; signal transduction; renal failure; nephritis;
KW hypothyroidism; hypogonadism; retinitis pigmentosa; growth disorder;
KW diabetes insipidus; hyperprolactinaemia; thirst disturbance; appetite;
KW sleep disturbance; temperature regulation; blood pressure; hypothalamus;
KW circadian rhythm.
XX
OS Mus sp.
XX
PH Key Location/Qualifiers
PT 1..1251
FT CDS /*tag= a
FT /product= "Mouse TGR346b protein"
XX
PN WO200242458-A2.
XX
PD 30-MAY-2002.
XX
XX 21-NOV-2001; 2001WO-US43404.
XX
XX 22-NOV-2000; 2000US-252841P.
XX 22-DEC-2000; 2000US-257636P.
XX 12-JAN-2001; 2001US-261377P.
XX 28-MAR-2001; 2001US-279554P.
XX 29-MAR-2001; 2001US-280696P.
XX
XX (TULA-) TULARIK INC.
XX
XX Tian H, Zhao J, Chen J, Cutler G, An S, Dai X, Gupte JS;
XX
XX WPI; 2002-463633/49.
XX P-PSDB; AAU97741.
XX
XX New isolated G-protein coupled receptor polypeptide, termed TGR, for
XX diagnosis and treatment of diseases such as renal failure, nephritis,
XX hypothyroidism, diabetes insipidus, and disturbances of thirst and
XX sleep
XX
XX Claim 22; Page 78-79; 98pp; English.
XX
XX The invention relates to a G-protein coupled receptor polypeptide (GPCR),
XX termed TGR, and its associated nucleic acid. The sequences of the
XX invention are useful for identifying a compound that modulates signal
XX transduction and for identifying a mammal having a TGR-associated
XX disorder. The proteins and nucleic acids are useful in diagnosis and
XX treatment of diseases or conditions such as renal failure, nephritis,
XX hypothyroidism, hypogonadism, retinitis pigmentosa, growth disorders,

CC diabetes insipidus, hyperprolactinaemia and disturbances of thirst,
CC sleep, temperature regulation, appetite, blood pressure or any other
CC syndrome or disease associated with the hypothalamus. The sequences can
CC be used in regulation of circadian rhythms, for use as genetic markers
CC for the identification of mutations associated with diseases resulting
CC from GPCR inactivation in particular cell types and for identification of
CC modulators of GPCR signal transduction. This sequence represents cDNA
CC encoding the mouse TGR346b polypeptide.
XX
SQ Sequence 1772 BP; 494 A; 390 C; 403 G; 485 T; 0 other;
Query Match 38.6%; Score 932; DB 24; Length 1772;
Best Local Similarity 82.0%; Pred. No. 1.2e-138;
Matches 1138; Conservative 0; Mismatches 235; Indels 15; Gaps 5;
Qy 359 GGGGCTTAAACATTACCCCGGAGAGTCTCTCGGCTGCTGGGACACACCTGAGCG 418
Db 3 GTCGTGGAACCTTGACCGCGGAGAGCTCTCGGGCTGCTTGGGTGACACACCTGAGCG 62
Qy 419 GGAGCAGTTTCATCGCTCTGTACCGGCTGCGACCGCTCGTCTACACCCAGAGCTGCCGGG 478
Db 63 CGCTCAGTTTCATCGGCACTATGGCTGCGGCACTGTGTCTCACCCGGGAGCTTCCGC 122
Qy 479 ACGGCCAAGCTGGCCCTCTGTCTCACCGGCTGCTCATCTTGGCCCTGGCGCTCTTTGG 538
Db 123 GCGCGCCAGGCTGGCCCTCTGTGTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 182
Qy 539 CAATGCTCTGGTGTTCACGTGTGACCCGACAGCCAGCCGATGCGGACCTGACCAACAT 598
Db 183 CAACGCCCTGGTGTGTATGTGTGACCCGACAGCCAGCCGATGCGGACCTGACCAACAT 242
Qy 599 CTTTATCTGCTCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATCCCGTCA 658
Db 243 CTTTATCTGCTCTTGGCGCTCAGTGACCTGCTCATCTTCTTCTGCAATCCCGTCA 302
Qy 659 CATGCTCCAGAAATTTCCGACAACTGGCTGGGGGGTCTTTTCAATTTGCAAGATGGTGCC 718
Db 303 CATGCTCCAGAAAGCTCTCGGACACCTGGCTGGGGGGTGGCTTCAATTTGCAAGATGGTGCC 362
Qy 719 ATTTGTCAGTCTACCGCTGTGTGACAGAAATCCTCAGTATGACCTGATGCTGTGTGGA 778
Db 363 ATTTGTCAGTCTACCGCTGTGTGACAGAAATCCTTACTATGACCTGCAATTTGTGGA 422
Qy 779 AAGGCACAGGACTTTGTGCTATCTTTTAAATGAAGTGGCAATACACCAACCGAAGGC 838
Db 423 AAGGCACAGGACTTTGTCCATCTTTTAAATGAAGTGGCAATACACCAATCAAAGC 482
Qy 839 TTTTCAATCTAGTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 898
Db 483 TTTTCAATCTAGTGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 542
Qy 899 CGTGCAACAACTTGAGATCAATATGACTTCTTATATGAAGAAAGAACACATCTGCTGTT 958
Db 543 TGTGCGAGCTTGAGATTAAGTATGACTTCTTATATGAAGAAAGAACACATCTGCTGCT 602
Qy 959 AGAAGAGTGGACCGCCCTGTGACCCAGAAAGATCTACACCACTTCACTCTTGTCACTCT 1018
Db 603 GGAAGAGTGGAGCAGCCCGCTGACCAAGAAATCTACACCACTTCACTCTTGTCACTCT 662
Qy 1019 CTTTCTCTCTCTTATGATGATGCTTATCTGTACAGTAAATTTGGTATGAACTTTG 1078
Db 663 CTTTCTCTCTCTTATGATGATGCTTATCTGTACAGTAAATTTGGTATGAACTTTG 722
Qy 1079 GATAAAGAAAGAGTTGGGGATGTTTCACTGCTTCTGCTTCTGCTTCTGCTTCTGCTTCT 1138
Db 723 GATCAAGAAAGAAATCGGGATGGCTCAGTCTCCGAACTATTTCATGGAAAGAAATGTT 782
Qy 1139 CAAATAGCCAGGAAAGAAAGAGAGTGTGATTTATGATGGTGCACAGTGGTGGTCTCTCT 1198
Db 783 CAAATAGCCAGGAAAGAAAGAGAGTGTGATCATGATGATGATGATGATGATGATGATG 842
Qy 1199 TGCTGTGTGCTGGGACCAATTCATGTTTCCATATGATGATGATGATGATGATGATGATG 1258

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 05:38:02 ; Search time 8744 Seconds
(without alignments)
11298.805 Million cell updates/sec

Title: US-10-070-241B-2
Perfect score: 2415
Sequence: 1 gccacggcgccaggaccct.....aaaaaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 2045481386 residues
Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rnd.*
36: em_htg_man.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	1668.4	69.1	1710	6	AX360197	AX360197 Sequence
2	1296	53.7	1296	6	AX360195	AX360195 Sequence
3	1296	53.7	1296	6	AX65928	AX65928 Sequence
4	1296	53.7	1296	6	BD095744	BD095744 Novel gua
5	1293	53.5	1293	6	BD012995	BD012995 Novel G p
6	1293	53.5	1293	6	BD014506	BD014506 Novel G p
7	1122.8	46.5	75164	9	AC093867	AC093867 Homo sapi
8	1122.8	46.5	141912	6	AX646791	AX646791 Sequence
9	1121.2	46.4	155347	2	AC005961	AC005961 Homo sapi
10	1078.4	44.7	1368	6	AX549224	AX549224 Sequence
11	1078.4	44.7	1368	9	AF411117	AF411117 Homo sapi
12	994.6	41.2	1737	6	AX665938	AX665938 Sequence
13	932	38.6	1772	6	AX665940	AX665940 Sequence
14	790.8	32.7	951	6	AX657434	AX657434 Sequence
15	693.4	28.7	170227	9	AC093816	AC093816 Homo sapi
16	309.2	12.8	227102	2	AC102844	AC102844 Mus muscu
17	309	12.8	432	6	AX277359	AX277359 Sequence
18	309	12.8	432	6	AX277395	AX277395 Sequence
19	309	12.8	432	6	AX354959	AX354959 Sequence
20	297	12.3	149538	2	AC110951	AC110951 Rattus no
21	297	12.3	249366	2	AC123096	AC123096 Rattus no
22	278.8	11.5	224761	9	AC026894	AC026894 Homo sapi
23	278.2	11.5	163126	9	AC124658	AC124658 Homo sapi
24	278.2	11.5	256164	9	HS400877	AJ400877 Homo sapi
25	273	11.3	170561	9	AP005482	AP005482 Homo sapi
26	273	11.3	193858	2	AC007734	AC007734 Homo sapi
27	273	11.3	199853	2	AP001077	AP001077 Homo sapi
28	270	11.2	168203	2	AP001902	AP001902 Homo sapi
29	269.8	11.2	6712	9	HS803419	AL82112 Homo sapi
30	269.8	11.2	135038	9	HUMYWXD703	L78810 Homo sapien
31	268.2	11.1	171582	9	AC021558	AC021558 Homo sapi
32	264.4	10.9	128117	9	AC004000	AC004000 Human PAC
33	263.6	10.9	165871	9	CNS01DVO	AL135998 Human chr
34	263.6	10.9	173803	9	CNS01DV2	AL13448 Human chr
35	263	10.9	109739	2	AC115437	Continuation (4 of
36	263	10.9	276163	2	AC130878	AC130878 Rattus no
37	261.4	10.8	16849	2	AC011219	AC011219 Homo sapi
38	260.2	10.8	168624	9	CNS01DTN	AL132801 Human chr
39	258.6	10.7	161504	9	AC006517	AC006517 Homo sapi
40	258	10.7	215580	2	AC140526	AC140526 Homo sapi
41	257.8	10.7	218485	2	AC127470	AC127470 Pan trogl
42	257.4	10.7	196703	2	AC096851	AC096851 Pan trogl
43	257	10.6	37000	9	AL663061	AL663061 Human DNA
44	256.4	10.6	44348	9	AC006046	AC006046 Homo sapi
45	256.2	10.6	174869	2	AP001840	AP001840 Homo sapi

ALIGNMENTS

RESULT 1
AX360197
LOCUS AX360197 1710 bp DNA linear PAT 13-FEB-2002
DEFINITION Sequence 3 from Patent WO0204518.
ACCESSION AX360197
VERSION AX360197.1 GI:18675761
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Bloomquist,B.T. and Zhelnin,L.
TITLE Human neuropeptide Y-like g protein-coupled receptor
JOURNAL Patent: WO 0204518-A 3 17-JAN-2002;

FEATURES
 source
 Bayer Corporation (US)
 Location/Qualifiers
 1..1710
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 419 a 439 c 429 g 423 t
 ORIGIN
 Query Match 69.1%; Score 1668.4; DB 6; Length 1710;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1669; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 GCACAGCGCCGACGACCCCTAGCTGGCGCTCCAGACCCGACAGCGTGGCGGCTCG 60
 DB 41 GCACAGCGCCGACGACCCCTAGCTGGCGCTCCAGACCCGACAGCGTGGCGGCTCG 100
 QY 61 CTTAGGGAAGAGCAAGGGAAGAACTTTATTGAAACCGGGAACATTTTGTGCTACTGAG 120
 DB 101 CTTAGGGAAGAGCAAGGGAAGAACTTTATTGAAACCGGGAACATTTTGTGCTACTGAG 160
 QY 121 ATGAGTCTCCAGTCTTTGGCTTCCGCTCTCTTTATGCTGGGTTGATCCCTGAGCTG 180
 DB 161 ATGAGTCTCCAGTCTTTGGCTTCCGCTCTCTTTATGCTGGGTTGATCCCTGAGCTG 220
 QY 181 CTCTCCTTTCCGGAACCTCCGCGGTGCAGCCTTAGAGCCCTCCGCGCGGCTGACTCCAG 240
 DB 221 CTCTCCTTTCCGGAACCTCCGCGGTGCAGCCTTAGAGCCCTCCGCGCGGCTGACTCCAG 280
 QY 241 AGTAGAGGAAGGAGCGGCTCCGCTGCTCCGGAAGCCCTGCTGCCCGCAGATG 300
 DB 281 AGTAGAGGAAGGAGCGGCTCCGCTGCTCCGGAAGCCCTGCTGCCCGCAGATG 340
 QY 301 CGGATGGCAGCGCAGTAGCGGCTGGCCCGCTCCGCGGAGCGACAGCAATGCGAG 360
 DB 341 CGGATGGCAGCGCAGTAGCGGCTGGCCCGCTCCGCGGAGCGACAGCAATGCGAG 400
 QY 361 CGTTTAACATTAACCGGAGCACTTCTCGGCTGCTGGGGAACCACTTGAACGCGGG 420
 DB 401 CGTTTAACATTAACCGGAGCACTTCTCGGCTGCTGGGGAACCACTTGAACGCGGG 460
 QY 421 AGCAGTTTCATCGCTGACCGCTGCGACCGCTGCTACACCCGAGAGCTCCGCGGAC 480
 DB 461 AGCAGTTTCATCGCTGACCGCTGCGACCGCTGCTACACCCGAGAGCTCCGCGGAC 520
 QY 481 GCGCCAGCTGGCCCTGCTGCTCACCGGCTGCTCATCTTCGCCCTGGGCTCTTTGGCA 540
 DB 521 GCGCCAGCTGGCCCTGCTGCTCACCGGCTGCTCATCTTCGCCCTGGGCTCTTTGGCA 580
 QY 541 ATGCTCTGCTGTTCTAGCTGGTGCACCGGAGGCGCATGCGACCGGTACCCACATCT 600
 DB 581 ATGCTCTGCTGTTCTAGCTGGTGCACCGGAGGCGCATGCGACCGGTACCCACATCT 640
 QY 601 TTATCTGCTCTTGGCGCTCAGTGACCTGCTCATACACCTTCTTCTGCTTCCGCTCACCA 660
 DB 641 TTATCTGCTCTTGGCGCTCAGTGACCTGCTCATACACCTTCTTCTGCTTCCGCTCACCA 700
 QY 661 TGCTCCGAACATTTCCGACAACTGGCTGGGGGTGCTTTCATTTGCAAGATGTTGCCAT 720
 DB 701 TGCTCCGAACATTTCCGACAACTGGCTGGGGGTGCTTTCATTTGCAAGATGTTGCCAT 760
 QY 721 TTGTCCAGTCTACCGCTGTTGTGACAGAAATCTCTACTATGACCTGCTTGTGGA 780
 DB 761 TTGTCCAGTCTACCGCTGTTGTGACAGAAATCTCTACTATGACCTGCTTGTGGA 820
 QY 781 GGCCACCGGGAATTTGCTATCTTTTAAATGAAGTGGCAATACCAACCGAAGGCTT 840
 DB 821 GGCCACCGGGAATTTGCTATCTTTTAAATGAAGTGGCAATACCAACCGAAGGCTT 880
 QY 841 TCACAAATGCTAGTGTGGTCTGGCTGGTGGCAGTATCGTAGATACCCATGTGCGACG 900
 DB 881 TCACAAATGCTAGTGTGGTCTGGCTGGTGGCAGTATCGTAGATACCCATGTGCGACG 940

QY 901 TGCACAACTTGAGATCAAAATATGACTTCTCTATATGAAAAGAAACACATCTGCTCTTAG 960
 DB 941 TGCACAACTTGAGATCAAAATATGACTTCTCTATATGAAAAGAAACACATCTGCTCTTAG 1000
 QY 961 AAGAGTGGACCAAGCCCTGTGCACCAAGAGATCTACACCACTTTCATCTTGTCACTCTCT 1020
 DB 1001 AAGAGTGGACCAAGCCCTGTGCACCAAGAGATCTACACCACTTTCATCTTGTCACTCTCT 1060
 QY 1021 TCCTCTGCTCTTATATGAGTATGCTTATCTCTATGAGTAAATTTGTTATGAACTTTGGA 1080
 DB 1061 TCCTCTGCTCTTATATGAGTATGCTTATCTCTATGAGTAAATTTGTTATGAACTTTGGA 1120
 QY 1081 TAAAGAAAAGAGTTGGGATGTTCAAGTCTTTCGAACTATTTCATGAAAAGAAATGTCCA 1140
 DB 1121 TAAAGAAAAGAGTTGGGATGTTCAAGTCTTTCGAACTATTTCATGAAAAGAAATGTCCA 1180
 QY 1141 AAATAGCCAGGAAGAAACAGAGCTGTCTATTATGATGCTGACAGTGGTGGCTCTTTTG 1200
 DB 1181 AAATAGCCAGGAAGAAACAGAGCTGTCTATTATGATGCTGACAGTGGTGGCTCTTTTG 1240
 QY 1201 CTGCTGCTGGSCACCATTCCTATGCTTCCATATGATGATGATGATGATGATGATGATGATG 1260
 DB 1241 CTGCTGCTGGSCACCATTCCTATGCTTCCATATGATGATGATGATGATGATGATGATG 1300
 QY 1261 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1320
 DB 1301 AGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1360
 QY 1321 CCACTCCATCTCTAATCCCATGCTGATGATGATGATGATGATGATGATGATGATGATG 1380
 DB 1361 CCACTCCATCTCTAATCCCATGCTGATGATGATGATGATGATGATGATGATGATGATG 1420
 QY 1381 TTTTGTCTGAGTTTGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1440
 DB 1421 TTTTGTCTGAGTTTGTATTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480
 QY 1441 GAATTCAGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
 DB 1481 GAATTCAGGAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1540
 QY 1501 TGGAGGAAACCAAGGAGAGCAATTCAGTATGATGATGATGATGATGATGATGATGATGATG 1560
 DB 1541 TGGAGGAAACCAAGGAGAGCAATTCAGTATGATGATGATGATGATGATGATGATGATGATG 1600
 QY 1561 AGACAGAGGAG 1620
 DB 1601 AGACAGAGGAG 1660
 QY 1621 AGAATTCCTCTTATGACAGTGGGCAATTAATTAATAAATATCTTCTAAT 1670
 DB 1661 AGAATTCCTCTTATGACAGTGGGCAATTAATTAATAAATATCTTCTAAT 1710

RESULT 2
 AX360195
 LOCUS AX360195 1296 bp DNA linear PAT 13-FEB-2002
 DEFINITION Sequence 1 from Patent WO0204518.
 ACCESSION AX360195
 VERSION AX360195.1 GI:18675760
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1
 REFERENCE
 Bloomquist, B.T. and Zhelmin, L.
 Human neurotrophin y-like g protein-coupled receptor
 Patent: WO 0204518-A 1 17-JAN-2002;
 Bayer Corporation (US)
 FEATURES
 Location/Qualifiers
 1..1296
 /organism="Homo sapiens"
 /mol_type="genomic DNA"

```
BASE COUNT      344 a   305 c   301 g   346 t
ORIGIN
Query Match      53.7%; Score 1296; DB 6; Length 1296;
Best Local Similarity 100.0%; Pred. No. 1.1e-254;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 354 ATGCAGGCGCTTAACATTAACCCCGAGCAGTTCTCTCGGCTGCTCGGGACACCAACCTG 413
DB 1 ATGCAGGCGCTTAACATTAACCCCGAGCAGTTCTCTCGGCTGCTCGGGACACCAACCTG 60
QY 414 ACGGGGAGCAGTTCATCGCTCTGTACCGGCTGCGACCGCTGCTGTACACCCAGAGTG 473
DB 61 ACGGGGAGCAGTTCATCGCTCTGTACCGGCTGCGACCGCTGCTGTACACCCAGAGTG 120
QY 474 CCGGACGCGCAAGCTGCGCCCTCGTGTCTACCGGCTGCTCATCTTCGCCCTGCGCGTCC 533
DB 121 CCGGACGCGCAAGCTGCGCCCTCGTGTCTACCGGCTGCTCATCTTCGCCCTGCGCGTCC 180
QY 534 TTTGGCAATGCTCTGTGTTTCTACGTGTGACCCGACGAGGCCATCGCACCCCTCACC 593
DB 181 TTTGGCAATGCTCTGTGTTTCTACGTGTGACCCGACGAGGCCATCGCACCCCTCACC 240
QY 594 AACATCTTTATCTGCTCTCGGCTCAGTGACCTGCTCATCACCCTTCTTCGCAATCCC 653
DB 241 AACATCTTTATCTGCTCTCGGCTCAGTGACCTGCTCATCACCCTTCTTCGCAATCCC 300
QY 654 GTCACCATGCTCCGAAACATTTCCGACAACTTGGCTGGGGGTCCTTTCATTGCAAGATG 713
DB 301 GTCACCATGCTCCGAAACATTTCCGACAACTTGGCTGGGGGTCCTTTCATTGCAAGATG 360
QY 714 GTGCCATTGTCCAGTCTACCGCTGTTGTGACAGAAATCTTCATCATGACCTGCAATTGCT 773
DB 361 GTGCCATTGTCCAGTCTACCGCTGTTGTGACAGAAATCTTCATCATGACCTGCAATTGCT 420
QY 774 GTGAAAGCGCACCGGAGCTTGTGCATCCTTTTAAATGAAAGTGGCAATACACCAACCGA 833
DB 421 GTGAAAGCGCACCGGAGCTTGTGCATCCTTTTAAATGAAAGTGGCAATACACCAACCGA 480
QY 834 AGGCGCTTTCACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 953
DB 481 AGGCGCTTTCACAACTGCTAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 894 TGGCAGTGTGCAACACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGTC 953
DB 541 TGGCAGTGTGCAACACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACATCTGTC 600
QY 954 TGCTTAGAAGAGTGGACCGCCCTGTGACCAAGAGATCTACACCACTTCACTCTGTC 1013
DB 601 TGCTTAGAAGAGTGGACCGCCCTGTGACCAAGAGATCTACACCACTTCACTCTGTC 660
QY 1014 ATCCTCTTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 1073
DB 661 ATCCTCTTCTCTGCTCTTATGCTGATGCTTATCTGTACAGTAAATTTGGTTATGAA 720
QY 1074 CTTTGGATAAGAAAGAGTTGGGATGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACT 1133
DB 721 CTTTGGATAAGAAAGAGTTGGGATGCTTCACTGCTTCACTGCTTCACTGCTTCACTGCTTCACT 780
QY 1134 ATGTCCAAATAGCAGGAAGAAAGAGCTGTGCTTATGATGCTGCTGCTGCTGCTGCTGCT 1193
DB 781 ATGTCCAAATAGCAGGAAGAAAGAGCTGTGCTTATGATGCTGCTGCTGCTGCTGCTGCT 840
QY 1194 CTCCTTGTGTGTGCTGGGACCACTTCCATGCTGTGCTATGATGATGATGATGATGATGATGATGAT 1253
DB 841 CTCCTTGTGTGTGCTGGGACCACTTCCATGCTGTGCTATGATGATGATGATGATGATGATGATGAT 900
QY 1254 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
DB 901 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1314 GGATTTTCCAACTCCATCTGTAATCCCACTGCTGCTATGCAATTAATGAATGAATGAATCTCAA 1373
```

Dd 61 ACGGGGAGAGTTTCATCGCTCTGTACCGGCTGGACCGCTCGTCTATACACCCAGAGCTG 120

Qy 474 CCGGAGCGCGCAAGCTGGCCCTCGTGGCTCACCGGGTGCTCATCTTCCGCTTGGGGTCC 533

Dd 121 CCGGAGCGCGCAAGCTGGCCCTCGTGGCTCACCGGGTGCTCATCTTCCGCTTGGGGTCC 180

Qy 534 TTGGCAATGCTCTGTGTGTTCTACGTGGTGAACCCGAGCAAGGCCATGGCACCGTCAAC 593

Dd 181 TTGGCAATGCTCTGTGTGTTCTACGTGGTGAACCCGAGCAAGGCCATGGCACCGTCAAC 240

Qy 594 AACATCTTTATCTGCTCTTGGGCTCAGTGAACCTGCTCATCACTTCTTTCGCAATTC 653

Dd 241 AACATCTTTATCTGCTCTTGGGCTCAGTGAACCTGCTCATCACTTCTTTCGCAATTC 300

Qy 654 GTCACCATGCTCCAGAACTTCCGAACTGCTGGGGGGTCTTCAATTTGCAAGATG 713

Dd 301 GTCACCATGCTCCAGAACTTCCGAACTGCTGGGGGGTCTTCAATTTGCAAGATG 360

Qy 714 GTGCCATTTGTCCAGTCTACCGGTGTTGTGACAGAAATCCTCACTATGACCTGCAATGCT 773

Dd 361 GTGCCATTTGTCCAGTCTACCGGTGTTGTGACAGAAATCCTCACTATGACCTGCAATGCT 420

Qy 774 GTGGAAGGCACAGAGACTTGTGCATCTCTTTAAATGAAGTGGCAATACCAACCGA 833

Dd 421 GTGGAAGGCACAGAGACTTGTGCATCTCTTTAAATGAAGTGGCAATACCAACCGA 480

Qy 834 AGGGCTTTTACAACTGCTAGTGTGCTGGCTGGTGGCAGTCACTGATGATCACCCTG 893

Dd 481 AGGGCTTTTACAACTGCTAGTGTGCTGGCTGGTGGCAGTCACTGATGATCACCCTG 540

Qy 894 TGGCAGTGCACAACTTGCAGTCAAAATGATGCTTCTTATGAAAGGAAACACATCTGC 953

Dd 541 TGGCAGTGCACAACTTGCAGTCAAAATGATGCTTCTTATGAAAGGAAACACATCTGC 600

Qy 954 TGCCTAGAGAGTGCACAGCCCTGTGCACGAGAGATCAGACACCTTCTCTTGTGTC 1013

Dd 601 TGCCTAGAGAGTGCACAGCCCTGTGCACGAGAGATCAGACACCTTCTCTTGTGTC 660

Qy 1014 ATCTCTTCTCTCTCTCTCTTATGTTGATGCTTATCTGACAGTAAATTTGTTATGAA 1073

Dd 661 ATCTCTTCTCTCTCTCTCTTATGTTGATGCTTATCTGACAGTAAATTTGTTATGAA 720

Qy 1074 CTTTGGATAAAGAAAGAGTTGGGATGTTTCAGTGTTCGAACTTATCATGGAAGAA 1133

Dd 721 CTTTGGATAAAGAAAGAGTTGGGATGTTTCAGTGTTCGAACTTATCATGGAAGAA 780

Qy 1134 ATGTCCAAATAGCAGGAAGAGAAACGAGCTGCATTTATGATGATGTTGATGCTGCT 1193

Dd 781 ATGTCCAAATAGCAGGAAGAGAAACGAGCTGCATTTATGATGATGTTGATGCTGCT 840

Qy 1194 CTCTTTGCTGTGCTGGSCACCATTTCCATGTTGTCATATGATGATGATGATGATGAT 1253

Dd 841 CTCTTTGCTGTGCTGGSCACCATTTCCATGTTGTCATATGATGATGATGATGATGAT 900

Qy 1254 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313

Dd 901 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

Qy 1314 GGATTTTCCATCTCTGTAATCCCATGTTGCTATGATTTATGATTTATGATTTATGAT 1373

Dd 961 GGATTTTCCATCTCTGTAATCCCATGTTGCTATGATTTATGATTTATGATTTATGAT 1020

Qy 1374 AAAAATGTTTGTCTGCAATTTGTTATGATAGTAAATAAAACCTTCTCTCCAGCAAA 1433

Dd 1021 AAAAATGTTTGTCTGCAATTTGTTATGATAGTAAATAAAACCTTCTCTCCAGCAAA 1080

Qy 1434 AGGCATGGAATTCAGGAATTCAGTATGCGGAAGAAAGCAAAATTTTCCCTCAGAGAG 1493

Dd 1081 AGGCATGGAATTCAGGAATTCAGTATGCGGAAGAAAGCAAAATTTTCCCTCAGAGAG 1140

Qy 1494 AATCCAGTGGAGAAACCAAGGAGAGAGTTCAGTATGAGCAAAATTTGAGTCAAAATG 1553

Dd 1141 AATCCAGTGGAGAAACCAAGGAGAGAGTTCAGTATGAGCAAAATTTGAGTCAAAATG 1200

Qy 1554 TGTGAACAGACAG 1613

Dd 1201 TGTGAACAGACAG 1260

Qy 1614 CTGGCTGAGAAATTCCTCTTTAGACAGTGGGCAATAA 1649

Dd 1261 CTGGCTGAGAAATTCCTCTTTAGACAGTGGGCAATAA 1296

RESULT 4
BD095744
LOCUS
DEFINITION
Novel guanosine triphosphate-bound protein-coupled receptors and
genes encoding them, and their production and use.
ACCESSION
BD095744
VERSION
WO 0148189-A/1 GI:22641332
KEYWORDS
WO 0148189-A/2.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1296)
AUTHORS
Matsumoto, S., Oda, T., Saito, Y., Noriyuki, Morikawa, Yoshida, K.,
Suwa, M. and Sugiyama, T.
TITLE
Novel guanosine triphosphate-bound protein-coupled receptors and
genes encoding them, and their production and use
JOURNAL
Patent: WO 0148189-A 2 05-JUL-2001;
HELIIX RESEARCH INSTITUTE, SHUNICHIRO MATSUMOTO, TAMAKI ODA, YOKO
SAITO, NORIYUKI MORIKAWA, KENJI YOSHIDA, MAKIKO SUWA, TOMOYASU
SUGIYAMA
COMMENT
OS Homo sapiens (human)
PN WO 0148189-A/2
PD 05-JUL-2001
PF 28-DEC-2000 WO 2000JP009409
PR 28-DEC-1999 JP 99P 375152.31-MAR-2000 JP 00P 101339 PR
23-MAY-2000 JP 00P 155978
PI SHUNICHIRO MATSUMOTO, TAMAKI ODA, YOKO SAITO, NORIYUKI PI
MORIKAWA, KENJI YOSHIDA,
PI MAKIKO SUWA, TOMOYASU SUGIYAMA
PC C12N15/09, C12N1/19, C12N1/21, C12N5/10, C07K14/705, PC
C07K16/28,
PC C12P21/02, C12Q1/02, C12Q1/68, A61K31/711, A61K48/00, A61P43/00, PC
G01N33/15,
PC G01N33/50
CC Novel guanosine triphosphate-bound protein-coupled receptors
and genes
CC encoding them, and their production and use
FH Key Location/Qualifiers
FT source 1. .1296
/organism='Homo sapiens (human)'.
FEATURES
source Location/Qualifiers
1. .1296
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
BASE COUNT 344 a 305 c 301 g 346 t
ORIGIN
Query Match 53.7%; Score 1296; DB 6; Length 1296;
Best Local Similarity 100.0%; Pred. No. 1.1e-254;
Matches 1296; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 354 ATGCAGGGCTTAACATTACCCGGAGCAGTCTCTCGGCTGTCGGGACCAACCTG 413

Dd 1 ATGCAGGGCTTAACATTACCCGGAGCAGTCTCTCGGCTGTCGGGACCAACCTG 60

Qy 414 ACGGGGAGAGTTCATCGCTCTGTCACCGGTGGACCGCTCGTCTACACCCAGAGCTG 473

Dd 61 ACGGGGAGAGTTCATCGCTCTGTCACCGGTGGACCGCTCGTCTACACCCAGAGCTG 120

Qy 474 CCGGGAGCGCAAGCTGCGCTCTGCTCACCGGGCTGCTCATCTTTCGCGCTCGCGCTC 533

```
Db 121 CCGGACGCGCCAGCTGCGCCCTCGTGTCTCACCGGGCTGCTCATCTTCGCCCTCGGGCTC 180
QY 534 TTTGGCAATGCTCTGTGTGTTCTACGFTGTGACCCGAGCAAGCCATGCGCACCGTCAAC 593
Db 181 TTTGGCAATGCTCTGTGTGTTCTACGFTGTGACCCGAGCAAGCCATGCGCACCGTCAAC 240
QY 594 AACATCTTTATCTGCTCTTGGGGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATGCC 653
Db 241 AACATCTTTATCTGCTCTTGGGGCTCAGTGACCTGCTCATCACCTTCTTCTGCAATGCC 300
QY 654 GTCAACATGCTCCAGAACATTTCCGACAACTGGTGGGGGTGCTTTCATTTGCAAGATG 713
Db 301 GTCAACATGCTCCAGAACATTTCCGACAACTGGTGGGGGTGCTTTCATTTGCAAGATG 360
QY 714 GTGCCATTTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATTTGCT 773
Db 361 GTGCCATTTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATTTGCT 420
QY 774 GTGGAAAGGACACGAGGACTTGTGCATCCTTTTAAATGAAGTGGCAATACACCAACCGA 833
Db 421 GTGGAAAGGACACGAGGACTTGTGCATCCTTTTAAATGAAGTGGCAATACACCAACCGA 480
QY 834 AGGGCTTTTACAAATGCTAGGTGCTGCTGGCTGGTGGAGTCACTCGTAGGATCACCCATG 893
Db 481 AGGGCTTTTACAAATGCTAGGTGCTGCTGGCTGGTGGAGTCACTCGTAGGATCACCCATG 540
QY 894 TGGCACTGTGCAACAACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACACATCTGC 953
Db 541 TGGCACTGTGCAACAACTTGAGATCAAAATATGACTTCTTATATGAAAGGAAACACATCTGC 600
QY 954 TGCTTAGAAGAGTGGACAGCCCTGTGCACAGAGATCTACACCACTTCACTCTTGTGCT 1013
Db 601 TGCTTAGAAGAGTGGACAGCCCTGTGCACAGAGATCTACACCACTTCACTCTTGTGCT 660
QY 1014 ATCCTCTTCTCCTCCTCTCTATGTGTGATGCTTATTTCTGACAGTAAATGTTATGAA 1073
Db 661 ATCCTCTTCTCCTCCTCTCTATGTGTGATGCTTATTTCTGACAGTAAATGTTATGAA 720
QY 1074 CTTTGGATAAGAAAGAGTTGGGATGTTGAGTGTGCTTTCAGTGTGCAATCATGAAAGAA 1133
Db 721 CTTTGGATAAGAAAGAGTTGGGATGTTGAGTGTGCTTTCAGTGTGCAATCATGAAAGAA 780
QY 1134 ATGTCAAAATAGCAGAGAGAAAGACGAGCTGTCATATGATGTCAGTGTGCT 1193
Db 781 ATGTCAAAATAGCAGAGAGAAAGACGAGCTGTCATATGATGTCAGTGTGCT 840
QY 1194 CTCCTTGTGTGTGCTGGGACCATTCCTATGTTGTCATATGATGATGATGATGATGATGAT 1253
Db 841 CTCCTTGTGTGTGCTGGGACCATTCCTATGTTGTCATATGATGATGATGATGATGATGAT 900
QY 1254 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313
Db 901 TTTGAAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
QY 1314 GGAATTTTCCAACTCCATCTGTAATCCCATTTGCTATGATGATGATGATGATGATGATGAT 1373
Db 961 GGAATTTTCCAACTCCATCTGTAATCCCATTTGCTATGATGATGATGATGATGATGATGAT 1020
QY 1374 AAAAATGTTTCTGCAAGTTGTTATTTGATGATGATGATGATGATGATGATGATGATGAT 1433
Db 1021 AAAAATGTTTCTGCAAGTTGTTATTTGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1434 AGGCAATGAAATTCAGGAATTTACAAATGATGCGGAAGAAAGAAAGTTTTCCTCAGAG 1493
Db 1081 AGGCAATGAAATTCAGGAATTTACAAATGATGCGGAAGAAAGAAAGTTTTCCTCAGAG 1140
QY 1494 AATCAGTGGAGGAACCAAGGAGAGCAATTCAGTGTGCAACATTTGATGATGATGATGAT 1553
Db 1141 AATCAGTGGAGGAACCAAGGAGAGCAATTCAGTGTGCAACATTTGATGATGATGATGAT 1200
QY 1554 TGTGAACAGACAGAGGAGAAAGAAAGTCAAAACGACATCTTGTCTCTTTAGTCTGAA 1613
Db 1201 TGTGAACAGACAGAGGAGAAAGAAAGTCAAAACGACATCTTGTCTCTTTAGTCTGAA 1260
```

```
QY 1614 CTGGCTGAGAAATTCCTTTTAGACAGTGGCATTA 1649
Db 1261 CTGGCTGAGAAATTCCTTTTAGACAGTGGCATTA 1296
```

RESULT 5

```
BD012995          1293 bp      DNA      linear      PAT 02-AUG-2002
LOCUS             Novel G protein coupled receptor protein and its DNA.
DEFINITION        BD012995
ACCESSION          BD012995
VERSION            BD012995.1  GI:22093184
KEYWORDS            WO 0116316-A/1.
SOURCE              Homo sapiens (human)
ORGANISM            Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE           1. (bases 1 to 1293)
AUTHORS             Watanabe,T., Kikuchi,K. and Shintani,Y.
TITLE               Novel G protein coupled receptor protein and its DNA
JOURNAL             Patent: WO 0116316-A 1 08-MAR-2001;
                    TAKEDA CHEMICAL INDUSTRIES LTD, TAKUYA WATANABE, KUNIKO KIKUCHI,
                    YASUSHI SHINTANI
COMMENT             OS Homo sapiens (human)
                    PN WO 0116316-A/1
                    PD 08-MAR-2001
                    PF 24-AUG-2000 WO 2000JP005684
                    PR 27-AUG-1999 JP 99P 241530
                    PI TAKUYA WATANABE, KUNIKO KIKUCHI, YASUSHI SHINTANI PC
                    C12N15/12, C12N15/09, C07K14/705, C07K16/28, C12P21/02, A61K45/00, PC
                    A61P43/00,
                    PC G01N33/566, G01N33/15, G01N33/50
                    CC
FEATURES             Location/Qualifiers
                    source          1..1293
                                /organism="Homo sapiens"
                                /mol_type="genomic DNA"
                                /db_xref="taxon:9606"
BASE COUNT          342 a 305 c 301 g 345 t
ORIGIN
```

Query Match 53.5%; Score 1293; DB 6; Length 1293;
Best Local Similarity 100.0%; Pred. No. 4.3e-254;
Matches 1293; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 354 ATGCAGGCGCTTAACATACCCCGAGAGTTCTCTGGCTGTGCGGACCAACCTG 413
Db 1 ATGCAGGCGCTTAACATACCCCGAGAGTTCTCTGGCTGTGCGGACCAACCTG 60
QY 414 ACGGGGAGAGTTTCATCGTCTCTACCGCTGCGAGCGTCTACCCACAGAGCTG 473
Db 61 ACGGGGAGAGTTTCATCGTCTCTACCGCTGCGAGCGTCTCTACCCACAGAGCTG 120
QY 474 CCGGGAACGCGCAAGCTGGCCCTGCTCTACCGGCTGTCTCATCTTCGCCCTGGCGTC 533
Db 121 CCGGGAACGCGCAAGCTGGCCCTGCTCTACCGGCTGTCTCATCTTCGCCCTGGCGTC 180
QY 534 TTTGGCAATGCTCTGTGTTCTAGCTGTGACCCGCAAGGCAATGCGCACCGTCAAC 593
Db 181 TTTGGCAATGCTCTGTGTTCTAGCTGTGACCCGCAAGGCAATGCGCACCGTCAAC 240
QY 594 AACATCTTTATCTGCTCTTGGGGCTCAGTGACCTGCTCATCCTTCTTCTGCAATCCC 653
Db 241 AACATCTTTATCTGCTCTTGGGGCTCAGTGACCTGCTCATCCTTCTTCTGCAATCCC 300
QY 654 GTCAACATGCTCCAGAACATTTCCGACAACTGGTGGGGGTGCTTTCATTTGCAAGATG 713
Db 301 GTCAACATGCTCCAGAACATTTCCGACAACTGGTGGGGGTGCTTTCATTTGCAAGATG 360
QY 714 GTGCCATTTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATTTGCT 773
Db 361 GTGCCATTTGTCAGTCTACCGTGTGTGACAGAAATCCTCACTATGACCTGCAATTTGCT 420
```


Db	541	TGGCAGCTGCAACAACCTTGAGATCAAATAATGACTTCCCTATATGAAGAAGAACACATCTGC	600
Qy	954	TGCTTAGAAGAGTGGACAGCCCTGTGCACCAGAGAGTCTACACCACTTCATCCTTGTC	1013
Db	601	TGCTTAGAAGAGTGGACAGCCCTGTGCACCAGAGAGTCTACACCACTTCATCCTTGTC	660
Qy	1014	ATCCTCTTCTCTCGCCCTTTATGGTGATGCTTTATTCTGTACAGTAATAATGGTTATGAA	1073
Db	661	ATCCTCTTCTCTCGCCCTTTATGGTGATGCTTTATTCTGTACAGTAATAATGGTTATGAA	720
Qy	1074	CTTTGGATTAAGAAAGAGTTGGGGATGGTTCACTGCTTCCAACATATTCATGGAAAAGAA	1133
Db	721	CTTTGGATTAAGAAAGAGTTGGGGATGGTTCACTGCTTCCAACATATTCATGGAAAAGAA	780
Qy	1134	ATGTCCAAAAATAGCAGGAAGAAACAGAGCTGTCAATTATGATGGTGACAGTGGTGGCT	1193
Db	781	ATGTCCAAAAATAGCAGGAAGAAACAGAGCTGTCAATTATGATGGTGACAGTGGTGGCT	840
Qy	1194	CTCTTTGCTGTGTGCTGGCACCAATTCCATGTTGTCATATGATGATTAACAGTAAT	1253
Db	841	CTCTTTGCTGTGTGCTGGCACCAATTCCATGTTGTCATATGATGATTAACAGTAAT	900
Qy	1254	TTTTGAAAGGAATATGATGATGTCCAAATCAAGATGATTTTGCTATCGTCAAAATTATT	1313
Db	901	TTTTGAAAGGAATATGATGATGTCCAAATCAAGATGATTTTGCTATCGTCAAAATTATT	960
Qy	1314	GGATTTTCCAACTCCATCTGTAATCCCATTGTCTATGCANTTATGAATGAAACTTCAAA	1373
Db	961	GGATTTTCCAACTCCATCTGTAATCCCATTGTCTATGCANTTATGAATGAAACTTCAAA	1020
Qy	1374	AAAAATGTTTGTCTGCAGTTTGTATTGTCATAGTAATAAAAACTTCTCTCCAGACAA	1433
Db	1021	AAAAATGTTTGTCTGCAGTTTGTATTGTCATAGTAATAAAAACTTCTCTCCAGACAA	1080
Qy	1434	AGCGNCGAAATTCAGGAATTAACATGATGGGAAGAAAGAAAGTTTCCCTCAGAGAG	1493
Db	1081	AGCGATGGAATTCAGGAATTAACATGATGGGAAGAAAGAAAGTTTCCCTCAGAGAG	1140
Qy	1494	AATCCAGTGGAGAAACAAAGGAGAGCATTCAGTGTATGGCAACATTGAAGTCRAATTG	1553
Db	1141	AATCCAGTGGAGAAACAAAGGAGAGCATTCAGTGTATGGCAACATTGAAGTCRAATTG	1200
Qy	1554	TGTGAACAGACAGAGGAGAAAGAAAGCTCAAACGACATCTGCTCTTTAGTCTGTAA	1613
Db	1201	TGTGAACAGACAGAGGAGAAAGAAAGCTCAAACGACATCTGCTCTTTAGTCTGTAA	1260
Qy	1614	CTGGCTCAGAAATTCCTCTTTAGACAGTGGGCAT	1646
Db	1261	CTGGCTCAGAAATTCCTCTTTAGACAGTGGGCAT	1293
<hr/>			
RESULT 7			
AC093867/C			
LOCUS	AC093867	75164 bp	DNA linear PRI 01-MAR-2002
DEFINITION	Homo sapiens BAC clone RP11-587P6 from 4, complete sequence.		
ACCESSION	AC093867	AC068256	
VERSION	AC093867.3	GI:15799661	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	1 (bases 1 to 75164)		
MEDLINE	Sulston,J.E. and Waterston,R.		
PUBMED	Toward a complete human genome sequence		
AUTHORS	Genome Res. 8 (11), 1097-1108 (1998)		
TITLE	2 (bases 1 to 75164)		
JOURNAL	Trani,L. and Meyer,R.		
REFERENCE	The sequence of Homo sapiens BAC clone RP11-587P6		
	Unpublished (2001)		
	3 (bases 1 to 75164)		

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Waterston,R.H.
Direct Submission
Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 75164)
Waterston,R.H.
Direct Submission
Submitted (28-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 75164)
Waterston,R.
Direct Submission
Submitted (01-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Sep 28, 2001 this sequence version replaced gi:15778786.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc>
Contact: sapiens@watson.wustl.edu
----- Summary Statistics
Center project name: H NH0587P06
Drafting Center: WIBR

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RPci-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P. Y., Zhao, B., Frengen, E., Taten, M., Catanese, J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-364P2, 2000 bp overlap.

Polymorphisms have been identified between AC093867 and AC093816.

The sequence of AC068256 has been incorporated into AC093867.

```

FEATURES
line sequence of AC068256 has been
Location/Qualifiers
1..75164
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-587P6"
/clone_lib="RPC1-11"
repeat region
995..1208

```


repeat_region	/rpt_family="HY3"	60057..60352	repeat_region	complement(139196..139483)	
repeat_region	/rpt_family="Alu"	63695..63963	repeat_region	/rpt_family="Alu"	
repeat_region	/rpt_family="Alu"	67457..67895	repeat_region	Query Match	46.4%; Score 1121.2; DB 2; Length 155347;
repeat_region	/rpt_family="MER4"	complement(69113..69482)	repeat_region	Best Local Similarity	99.0%; Pred.No. 1.9e-218;
repeat_region	/rpt_family="THE1"	complement(69879..70176)	repeat_region	Matches 1126; Conservative	0; Mismatches 11; Indels 0; Gaps 0;
repeat_region	/rpt_family="Alu"	70985..71138	Qy	1249	GTAATTTTGAAGGAATATGATGATGCATCAATCAAGATGATTTTGTCTATCGTCAAA 1308
repeat_region	/rpt_family="LTR8"	71150..71472	Db	59887	GTAATTTTGAAGGAATATGATGATGCATCAATCAAGATGATTTTGTCTATCGTCAAA 59946
repeat_region	/rpt_family="THE1"	71478..71963	Qy	1309	TTATGGATTTTCAACTCCATCTGTAATCCCATGTCATGCAATTTATGAATGAAACT 1368
repeat_region	/rpt_family="LTR8"	complement(72510..72712)	Db	59947	TTATGGATTTTCAACTCCATCTGTAATCCCATGTCATGCAATTTATGAATGAAACT 60006
repeat_region	/rpt_family="MER21"	76772..77064	Qy	1369	TCAAAAAAATGTTTGTCTGCAATTTGTTATTCATAGTAAATAAAACCTTCTCTCAG 1428
repeat_region	/rpt_family="Alu"	82246..82490	Db	60007	TCAAAAAAATGTTTGTCTGCAATTTGTTATTCATAGTAAATAAAACCTTCTCTCAG 60066
repeat_region	/rpt_family="Alu"	complement(86124..86418)	Qy	1429	CACAAAGGCATGGAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGATTTTCCCTCA 1488
repeat_region	/rpt_family="Alu"	87429..87692	Db	60067	CACAAAGGCATGGAATTCAGGAATTAACAATGATGCGGAAGAAAGCAAGATTTTCCCTCA 60126
repeat_region	/rpt_family="Alu"	92523..92927	Qy	1489	GAGGAATCCAGTGGAGAAACCAAGAGAGCAATTCAGTATGCGCAATTCAGATCA 1548
repeat_region	/rpt_family="Alu"	92932..93218	Db	60127	GAGGAATCCAGTGGAGAAACCAAGAGAGCAATTCAGTATGCGCAATTCAGATCA 60186
repeat_region	/rpt_family="Alu"	98713..98971	Qy	1549	AATTGTGTGAACAGACAGAGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGT 1608
repeat_region	/rpt_family="LFR5"	complement(101354..101559)	Db	60187	AATTGTGTGAACAGACAGAGAGAGAAAGCTCAAAACGACATCTTGTCTCTTTAGGT 60246
repeat_region	/rpt_family="Alu"	complement(101576..101683)	Qy	1609	CTGAATGGCTGAGAAATTCCTTTAGACAGTGGGCATTAATTAACAATATCTTCA 1668
repeat_region	/rpt_family="Alu"	102229..102552	Db	60247	CTGAATGGCTGAGAAATTCCTTTAGACAGTGGGCATTAATTAACAATATCTTCA 60306
repeat_region	/rpt_family="LFR5"	102558..102665	Qy	1669	ATTAATGCCCTTCAGATTTACCCAAAGAGAAATTAATTTGAGCAAGATCAATCT 1728
repeat_region	/rpt_family="LFR5"	complement(104177..104560)	Db	60307	ATTAATGCCCTTCAGATTTACCCAAAGAGAAATTAATTTGAGCAAGATCAATCT 60366
repeat_region	/rpt_family="L1"	complement(105291..110361)	Qy	1729	CTTTTAATTTCTTAAGATGATGACAAGAGAAACAAATCATGTTTCCATTAAAAAATGAC 1788
repeat_region	/rpt_family="L1"	complement(108878..109028)	Db	60367	CTTTTAATTTCTTAAGATGATGACAAGAGAAACAAATCATGTTTCCATTAAAAAATGAC 60426
repeat_region	/rpt_family="MER25"	complement(110483..110652)	Qy	1789	ACGAGGCTAGTCCAAAGTGCAGTATGTTTACAAACCAATNGATCAATATTAAACANAT 1848
repeat_region	/rpt_family="L1"	complement(112817..113106)	Db	60427	ACGAGGCTAGTCCAAAGTGCAGTATGTTTACAAACCAATNGATCAATATTAAACANAT 60486
repeat_region	/rpt_family="Alu"	complement(113723..113889)	Qy	1849	TTCTGTGTTCTTCTCATTTCCACTGCTTCACTTAGCTAGCTTTAAAAAAGCAATGGA 1908
repeat_region	/rpt_family="Alu"	complement(116719..116785)	Db	60487	TTCTGTGTTCTTCTCATTTCCACTGCTTCACTTAGCTAGCTTTAAAAAAGCAATGGA 60546
repeat_region	/rpt_family="L1"	complement(120455..120511)	Qy	1909	AGGCCAGGCACGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCTAGACGGGGGG 1968
repeat_region	/rpt_family="Alu"	complement(123993..124318)	Db	60547	AGGCCAGGCACGGTGGCTCATGCTCTTAATCCAGCACTTTGGGAGGCTAGACGGGGGG 60606
repeat_region	/rpt_family="Alu"	complement(124910..125066)	Qy	1969	ATCACGAGGTCAGGAGATCAAAACCATCTGGCTTAACACGGTGAACCCCACTCTGTCT 2028
repeat_region	/rpt_family="L1"	complement(129701..130088)	Db	60607	ATCACGAGGTCAGGAGATCAAAACCATCTGGCTTAACACGGTGAACCCCACTCTGTCT 60666
repeat_region	/rpt_family="Alu"	complement(129701..130088)	Qy	2029	AAAATACAAAATTTAGCGGGCGTGGTGGGGGCACTGTAGTCCAGCTACTTTGGGAGC 2088
repeat_region	/rpt_family="Alu"	complement(129701..130088)	Db	60667	AAAATACAAAATTTAGCGGGCGTGGTGGGGGCACTGTAGTCCAGCTACTTTGGGAGC 60726
repeat_region	/rpt_family="Alu"	complement(129701..130088)	Qy	2089	CTCAGGGGGGAGATGGTGTGAACCCCGGAGCGGAGCTTGCAGTGTCCAGATCATGTC 2148
repeat_region	/rpt_family="Alu"	complement(129701..130088)	Db	60727	CTCAGGGGGGAGATGGTGTGAACCCCGGAGCGGAGCTTGCAGTGTCCAGATCATGTC 60786
repeat_region	/rpt_family="L1"	complement(134127..135016)	Qy	2149	CACCTGCATCCAGCTGGCGAAAGAGCGAGACTCCCGCTCAAAAATAATTTTTTTGA 2208
repeat_region	/rpt_family="L1"	complement(137132..138120)	Db	60787	CACCTGCATCCAGCTGGCGAAAGAGCGAGACTCCCGCTCAAAAATAATTTTTTTGA 60846
repeat_region	/rpt_family="LFR8"	complement(137269..137558)	Qy	2209	AAAATTCGTAAACCATCTTTTAAGATTTATTCAGTGAATTTTAAATAATCTTGTACAGA 2268

Db 60847 AAAATTCGTAACCATACATCTTTTAAAGATATATTCAGTGGATTTTAAAAATCTTGTAACAGA 60906

Qy 2269 AATCAGGGTCTTAGCTAGCAGATTTTCTCCACGCGAGTCACTGTAAATGTGACTATGTAT 2328

Db 60907 AATCAGGGTCTTAGCTAGCAGATTTTCTCCACGCGAGTCACTGTAAATGTGACTATGTAT 60966

Qy 2329 TGCTAGATTGAATAAGAAAATAAAATAATATCTTCTTCCCTTGNAAAAAATAAAAAA 2385

Db 60967 TGCTAGATTGAATAAGAAAATAAAATAATATCTTCTTCCCTTGNAAAAAATAAAAAA 61023

RESULT 10

AX549224

LOCUS AX549224 1368 bp DNA linear PAT 26-NOV-2002

DEFINITION Sequence 509 from Patent WO02061087.

ACCESSION AX549224

VERSION AX549224.1 GI:25813923

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Burmer, G.C., Roush, C.L. and Brown, J.P.

TITLE Antigenic peptides, such as for G protein-coupled receptors (GPCRs), antibodies thereto, and systems for identifying such antigenic peptides

JOURNAL Patent: WO 02061087-A 509 08-AUG-2002;

FEATURES

Location/Qualifiers

1..1368

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

BASE COUNT 345 a 346 c 319 g 358 t

ORIGIN

Query Match 44.7%; Score 1078.4; DB 6; Length 1368;

Best Local Similarity 91.3%; Pred. No. 3.6e-210;

Matches 1213; Conservative 0; Mismatches 1; Indels 114; Gaps 2;

Qy 322 CGGTGGCCCGCTCCGGAGGCGCACAGCAATGAGGCGCTTAACATTAACCCGAGC 391

Db 155 CGGTGGCCCGCTCCGGAGGCGCACAGCAATGAGGCGCTTAACATTAACCCGAGC 214

Qy 382 AGTTCTCTCGCTGCTCGGGACCAACCTGACCGGGAGCAGTTTCATCGCTCTGTACC 441

Db 215 AGTTCTCTCGCTGCTCGGGACCAACCTGACCGGGAGCAGTTTCATCGCTCTGTACC 274

Qy 442 GGCTGGACCGCTGCTCTACACCCAGAGCTGCGGGAGCGGCCAAGCTGGCCCTGTGC 501

Db 275 GGCTGGACCGCTGCTCTACACCCAGAGCTGCGGGAGCGGCCAAGCTGGCCCTGTGC 334

Qy 502 TCACCGGCGTCTCATCTTCCGCTGCGCTCTTGGCAATGCTCGTGTCTTACGTGG 561

Db 335 TCACCGGCGTCTCATCTTCCGCTGCGCTCTTGGCAATGCTCGTGTCTTACGTGG 394

Qy 562 TGACCCGAGCAAGCCATGCGACCGCTACCAACATCTTTATCTGCTTGGCGCTCA 621

Db 395 TGACCCGAGCAAGCCATGCGACCGCTACCAACATCTTTATCTGCTTGGCGCTCA 454

Qy 622 GTGACCTGCTCATCATCTTCTGTCATTCGGTCAACATGCTCCGTCAGAAATTTCCGACA 681

Db 455 GTGACCTGCTCATCATCTTCTGTCATTCGGTCAACATGCTCCGTCAGAAATTTCCGACA 514

Qy 682 ACTGCTGGGGGCTCTTCAATTTGCAAGATGGTGGCAATTTGTCCAGTCTACCGCTGTG 741

Db 515 ACTGCTGGGGGCTCTTCAATTTGCAAGATGGTGGCAATTTGTCCAGTCTACCGCTGTG 574

Qy 742 TGACAGAAATCTCTATGACCTGTGCTGTGGAAAGGCAACAGGACTTGTGCATC 801

Db 575 TGACAGAAATCTCTATGACCTGTGCTGTGGAAAGGCAACAGGACTTGTGCATC 634

Qy 802 CTTTAAAAATGAAGTGGCAATACACCAACCGAAGGCGCTTTCACATGCTAGTGTGGTCT 861

Db 635 CTTTAAAAATGAAGTGGCAATACACCAACCGAAGGCGCTTTCACATGCTAGTGTGGTCT 694

Qy 862 GGCTGGTGGCAGTCACTCGTAGGATCACCCATGTCGACGTCGCAACAACTTGAGATCAAT 921

Db 695 GGCTGGTGGCAGTCACTCGTAGGATCACCCATGTCGACGTCGCAACAACTTGAGATCAAT 754

Qy 922 ATGACTTCTTATATGAAGAAACACATCTGCTGTCTTAGAAGAGTGGACAGCCCTGTGC 981

Db 755 ATGACTTCTTATATGAAGAAACACATCTGCTGTCTTAGAAGAGTGGACAGCCCTGTGC 814

Qy 982 ACCAGAGATCTACACCACTTCTATCTTGTCTATCTCTCTCTCTCTCTCTCTCTCTCT 1041

Db 815 ACCAGAGATCTACACCACTTCTATCTTGTCTATCTCTCTCTCTCTCTCTCTCTCTCT 868

Qy 1042 TGTCTTATCTGTACAGTAAATTTGGTTATGAACTTTGGATAAAGAAAGATTTGGGGATG 1101

Db 869 ----- 868

Qy 1102 GTTCAGTGTCTCGAATATTCATGGAAGAAATGTCAAAATAGCAGCAAGAAAGAAC 1161

Db 869 -----GGAAGAAAGAAC 880

Qy 1162 GAGCTGTCAATATCATGTCGACAGTGGTCTCTCTTGTCTGTCTGTCTGTCTGTCTGTCT 1221

Db 881 GAGCTGTCAATATCATGTCGACAGTGGTGGTCTCTCTTGTCTGTCTGTCTGTCTGTCTGT 940

Qy 1222 ATGTTGTCATATGATTAATGAAATACAGTAAATTTTGAAGAGATATGATGTACAA 1281

Db 941 ATGTTGTCATATGATTAATGAAATACAGTAAATTTTGAAGAGATATGATGTACAA 1000

Qy 1282 TCAAGATGATTTTGTCTATGTCGCAATTTTGGATTTTCCAACTCCATCTGTAATCCCA 1341

Db 1001 TCAAGATGATTTTGTCTATGTCGCAATTTTGGATTTTCCAACTCCATCTGTAATCCCA 1060

Qy 1342 TTGCTATGATTTTATGAATGAAATCTCAAAAAAATGTTTGTCTGCGATTTGTTATT 1401

Db 1061 TTGCTATGATTTTATGAATGAAATCTCAAAAAAATGTTTGTCTGCGATTTGTTATT 1120

Qy 1402 GCATAGTAAATAAAACCTTCTCCAGCACAAGGCGATGGAATTCAGGAATTACATGA 1461

Db 1121 GCATAGTAAATAAAACCTTCTCCAGCACAAGGCGATGGAATTCAGGAATTACATGA 1180

Qy 1462 TGCGAAGAAAGCAAAAGTTTCCCTCAGAGAGATTCAGTGGAGAAACCAAGGAGAAG 1521

Db 1181 TGCGAAGAAAGCAAAAGTTTCCCTCAGAGAGATTCAGTGGAGAAACCAAGGAGAAG 1240

Qy 1522 CATTCAGTGTGGCAACATGAGTCAATTTGTGTAACAGACAGAGAGAGAAAGAGC 1581

Db 1241 CATTCAGTGTGGCAACATGAGTCAATTTGTGTAACAGACAGAGAGAGAAAGAGC 1300

Qy 1582 TCAACAGACATCTTGTCTCTTTAGGCTGAACTCGGCTGAGAAATTCCTTTAGACAGTG 1641

Db 1301 TCAACAGACATCTTGTCTCTTTAGGCTGAACTCGGCTGAGAAATTCCTTTAGACAGTG 1360

Qy 1642 GGCATTAA 1649

Db 1361 GGCATTAA 1368

RESULT 11

AF411117

LOCUS AF411117 1368 bp mRNA linear PRI 01-NOV-2001

DEFINITION Homo sapiens G protein-coupled receptor (GPR103) mRNA, complete cds.

ACCESSION AF411117

VERSION AF411117.1 GI:16566346

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Gupte,J.S.
G-protein coupled receptors
Patent: WO 0242458-A 17 30-MAY-2002;
Tularik Inc. (US)
FEATURES
Location/Qualifiers
1..1772
/organism="Mus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10095"
1..1251
/note="unnamed protein product; mouse G-protein coupled receptor (GPCR) TGR346b"
/codon_start=1
/protein_id="CAD82895.1"
/db_xref="GI:29290814"
/translation="MSWNLTAQLSALLRLHNLRAQTAHVLGLRLPLVLTLPOLPARAR
LALLLVGLMIIFALALFGLNALVVVTVTRSKAMETVNIIFCSLALSDLLIVFECIPVTM
LQNVSDTWLGGAFICKMPFVOCATVTELLTMTCTIAVERHGLVHPFKMKQYINQR
AFTMLGVWLVAIIIGSPMWHVORLEIKDFLYEKHEHCCLLEWSPVHQKIYTFIL
VTLFLPLLLLSLYKIGYELWIKKRIGDGSVLRTHGKEMFKIARKKRAVIMVT
VVLFPAVCWAPPHIVHMMIETPSNFKEYDEVTKMIFAIQVIGIFNFCINPIIYALM
NENFKNFVSVCYICVKEETPSARKHSGSGAMVHRRAKLAARENPEIKGEAFGS
NIDIKWCQPEKKRKRSVASPL"
BASE COUNT 494 a 390 c 403 g 485 t
ORIGIN
Query Match 38.6%; Score 932; DB 6; Length 1772;
Best Local Similarity 82.0%; Pred. No. 3.4e-180;
Matches 1138; Conservative 0; Mismatches 235; Indels 15; Gaps 5;
QY 359 GCGCTTAACATTACCCGGACAGTTCTCTCGGCTGCTCGGGACCAACACTGACGGG 418
Db 3 GTCTGGGAACCTTGACCGGGAGCAGCTCTCGGCTGCTCGGCTGCAACCTGACGGG 62
QY 419 GGAGCAGTTTCATCGCTCTGTACCGGCTGGACCGCTGCTGTCTACACCCAGAGCTCGCGG 478
Db 63 CGTCAGTTTCATCGGCACTATGGGCTGGGCACTGGTGTCTACCCGCACTTCCCGC 122
QY 479 AGCGCCAAAGTGGCCCTGCTGCTCACCGGCTGCTCATCTTCGCGCTGGCGCTCTTTGG 538
Db 123 GCGGCCAGGCTGGCCCTCTGCTGGTGGGCTGCTGCTCTTTTGGCGCTCTTGG 182
QY 539 CAATGCTCTGTGTTCTAGTGTGACCGGACGAGCGCATGGCAGCGTCAACCAAT 598
Db 183 CAACGCCCTGTGTGCTATGTGTGACCGGACGAGCGCATGGCAGCGTCAACCAAT 242
QY 599 CTTTATCTGCTCTTGGGCTCAGTGACCTGCTCATCTCTCTCTCTCTCTCTCTCTCTCT 658
Db 243 CTTTATCTGCTCTTGGGCTCAGTGACCTGCTCATCTCTCTCTCTCTCTCTCTCTCTCT 302
QY 659 CATGTCAGAACATTTCCGAACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGCC 718
Db 303 CATGTCAGAACATTTCCGAACTGGCTGGGGGTGCTTTTCATTTGCAAGATGGTGCC 362
QY 719 ATTGTCCAGTCTACCGTGTGTGACAGAAATCTCTACTATGACTGCTGCTGTGGA 778
Db 363 ATTGTCCAGTCTACCGTGTGTGACAGAAATCTCTACTATGACTGCTGCTGTGGA 422
QY 779 AAGGACACGAGGACTTGTGCAATCTTTTAAATGAAGTGGCAATACACCAACCGAAGGC 838
Db 423 AAGGACACGAGGACTTGTGCAATCTTTTAAATGAAGTGGCAATACCAATCAAGAGC 482
QY 839 TTTCAATGT 898
Db 483 TTTCAATGT 542
QY 899 CGTGCAACAACTTGAGATCAATATGACTTCTATATGAAAGGAACCACTGTGTGTT 958
Db 543 TGTGAGGAGCTTGAGATTAAGTATGACTTCTATATGAAAGGAACCACTGTGTGCT 602
QY 959 AGAAGGTGGACCGCTGTGACCAAGAGATCTACACCGCTTCTCTCTCTCTCTCTCTCT 1018
Db 603 GGAAGGTGGACCGCTGTGACCAAGATCTACACCGCTTCTCTCTCTCTCTCTCTCTCT 662

QY 1019 CTTCTCTGCTCTTATGTGTGATGCTTATTTCTGTACAGTAAATTTGGTATGAACTTTG 1078
Db 663 CTTCTCTGCTACACTGTTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 722
QY 1079 GATAAAGAAAAGAGTTGGGGATGGTTCAAGTCTTCGAACTATTTCATGGAAAAGAAATGTC 1138
Db 723 GATCAAGAAAAGAAATCGGGATGGCTCAGTCTCTCGAACTATTTCATGGAAAAGAAATGTT 782
QY 1139 CAATAATAGCCAGGAGAGAGAAACGAGCTGTCTATTATGATGGTGACAGTGGTGGCTCTCT 1198
Db 783 CAATAATAGCCAGGAGAGAGAAACGAGCTGTCTATTATGATGGTGACAGTGGTGGCTCTCT 842
QY 1199 TGCTGTGTGTGTGGGACCATTTCCATTTGTTGTTCCATATGATGATTGAATACAGTAATTTTGA 1258
Db 843 TGCTGTGTGTGTGGGACCATTTCCACATCGTTCACATGATGATTGAATACAGTAATTTTGA 902
QY 1259 AAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1318
Db 903 AAAGGAATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 962
QY 1319 TTCAAACCTCCATCTGTAATCCATGCTGCTATGCTATGCTATGCTATGCTATGCTATGCT 1378
Db 963 TTTCACTCCATCTGTAATCCATTTATTTATGCACTTATGATGATAAATACTTTCAAAAAA 1022
QY 1379 TGTTTGTCTGAGTTTGTATTGCTAGTAGTAATAAATAAATACTTTCTCCAGCAACAAGCA 1438
Db 1023 CTTTGTGTCTGCGGTTTGTCTATTGCTATTGCTAGGAGAAACACTTTCTTCCAGCAGGA 1082
QY 1439 TGGAAATTCAGGAATTAACAATGATGCGGAGAGAAAGCAAGATTTTCCCTCAGAGAGATCC 1498
Db 1083 TGGAAATTCAGGAGTATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1142
QY 1499 AGTGAGAGAACCAAGAGAGAGCAATTCAGTGATGCGAAGCAATTTGAAGTCAATTTGTTGA 1558
Db 1143 TGTAGAG---ATCAAGAGAGAGCAATTTGGGGGAGCAACATCGATATCAAGTGTGTGA 1199
QY 1559 ACAGCAGAGAGAGAGAGAA---AAAGCTCAACAGCACTTTGCTCTCTTTAGTCTGTAAC 1614
Db 1200 ACAGCAGAGAGAGAGAGAGAGAGATCAAAAGTGGCATCTTTGCTCTTTAGTTCGAA- 1258
QY 1615 TGGCTGAGAAATTCCTTTAGACAGTGGGCAATTAATATAACAATATCTTCATAATTAAT 1674
Db 1259 TTTCTGAGAGCTCTGCTGTAGAC-----GTCAACACTGTACCAATGTCTTCAGAAATGAGT 1313
QY 1675 GCCCTTCAGATTTAAACCAAGAGAGAAATTTATTTGAGCAAGAGTCAAACTCTTTTA 1734
Db 1314 ATCTGTATATCTGTAAATCGAA--AGAAAATGATTTTGAGAAAAAGCCAGAGAGCTTTTCA 1371
QY 1735 ATTCTTAA 1742
Db 1372 ATTAAGAA 1379
RESULT 14
AX657434
LOCUS 951 bp DNA linear PAT 22-MAR-2003
DEFINITION Sequence 17 from Patent WO03000893.
ACCESSION AX657434
VERSION AX657434.1 GI:29160173
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1
AUTHORS Martinez R.M. and Sigurdsson G.T.
TITLE Nucleic acids encoding G protein-coupled receptors
JOURNAL Patent: WO 03000893-A 17 03-JAN-2003;
Decode Genetics EHF. (IS)
FEATURES
Location/Qualifiers
1..951
/organism="Homo sapiens"
source

```

/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>951
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAD79723.1"
/db_xref="GI:29160174"
/translation="PGRKMLVLTGLVLFALFGLNALVFFVTVRSKTMRTVNIIFI
CSLASDMLITLFCIPVTMLQNSDMLGVAFCIKMPFVQSTAVVTELTTCIAVE
RHQGLVHPKPMKQVNTNRAFTMLGVVLAVIVGSPMWHIKYDFLYEKEHLCLEEM
TSPVHXKIYTTILATILFLPLMWMLILYSKIGYELMIKKRKRAVIMVTVVAFV
CHAPFHVHMVSEBFLSHVTPSMFAIVQIIGFSNCFINFIYAFMNEFKKNVLS
AVCICIVNKTSPAPRHGNSGITMRKAKFSLRNPVBET"
BASE COUNT      241 a    231 c    206 g    273 t
ORIGIN

Query Match      32.7%; Score 790.8; DB 6; Length 951;
Best Local Similarity 89.0%; Pred. No. 2.3e-151;
Matches 924; Conservative 0; Mismatches 27; Indels 87; Gaps 3;

QY 474 CCGGACGGCCCAAGCTGGCCCTCGTCTCACCGGGTGTCTATCTTCGCCCTGGCGTC 533
Db      |||||
QY 1 CCGGACGGCCCAAGATGCGCCCTCGTCTCACCGGGTGTCTATCTTCGCCCTGGCGTC 60
QY 534 TTTGGCAATGCTCTGTGTCTTACGTGTGTGACCCGAGCAAGCCATGCGACCGTCACC 593
Db      |||||
QY 61 TTTGGCAATGCTCTGTGTCTTACGTGTGTGACCCGAGCAAGCCATGCGACCGTCACC 120
QY 594 AACATCTTTATCTGCTCTGCGCTCAGTGCACCTCATCATCCTTCTTCTGCATTCCC 653
Db      |||||
QY 121 AACATCTTTATGCTCTGCGCTCAGTGCACCTCATCATCCTTCTTCTGCATTCCC 180
QY 654 GTCCACATGCTCCAGAACATTTCCGAACTGGCTGGGGGGTGTCTTCAATTGCAAGATG 713
Db      |||||
QY 181 GTCCACATGCTCCAGAACATTTCCGAACTGGCTGGGGGGTGTCTTCAATTGCAAGATG 240
QY 714 GTGCCATTTGTCCAGTCTACCGCTGTGTGACAGAAATCCTCACTATGACCTGCATGCT 773
Db      |||||
QY 241 GTGCCATTTGTCCAGTCTACCGCTGTGTGACAGAAATCCTCACTATGACCTGCATGCT 300
QY 774 GTGAAAGGCACACAGGACTTGTGCATCTTTTAAATGAAGTGCATACACCAACCGA 833
Db      |||||
QY 301 GTGAAAGGCACACAGGACTTGTGCATCTTTTAAATGAAGTGCATACACCAACCGA 360
QY 834 AGGGCTTTACAAATGCTAGGTGTGTCTGGCTGTGGGAGTCACTGATGATCAACCCATG 893
Db      |||||
QY 361 AGGGCTTTACAAATGCTAGGTGTGTCTGGCTGTGGGAGTCACTGATGATCAACCCATG 420
QY 894 TGGCAGTGCACAACTTGAGATCAATATGACTTCTTATATGAAAGAACACATCTGC 953
Db      |||||
QY 421 TGGCAC-----ATCAATATGACTTCTTATATGAAAGAACACATCTGC 465
QY 954 TGTCTAGAGAGTGGACAGCCCTGTGCACAGAAAGATCTACACCACTTCATCTTGTGTC 1013
Db      |||||
QY 466 TGTCTAGAGAGTGGACAGCCCTGTGCACAGAAAGATCTACACCACTTCATCTTGTGTC 525
QY 1014 ATCCTCTTCTCTCGCTCTTATGATGATGCTTATCTGTACAGTAAATGTTGTTATGAA 1073
Db      |||||
QY 526 ATCCTCTTCTCTCGCTCTTATGATGATGCTTATCTGTACAGTAAATGTTGTTATGAA 585
QY 1074 CTTTGGATAAAGAAAGAGTTGGGATGTTTCAGTGTCTCGAACTATTCATGGAAGAA 1133
Db      |||||
QY 586 CTTTGGATAAAGAAAGAA----- 603
QY 1134 ATGTCCAAATAGCCAGGAAGAAAGAAAGAGTGTCAATTATGATGTCAGTGTGCT 1193
Db      |||||
QY 604 -----AAGAAACGAGTGTCAATTATGATGTCAGTGTGCT 642
QY 1194 CTCCTTGTGTGTGGCCACCATTCACATGTTGTCATATGATGATGATGATGATGATGAT 1253
Db      |||||
QY 643 CTCCTTGTGTGTGGCCACCATTCACATGTTGTCATATGATGATGATGATGATGATGAT 702
QY 1254 TTTGAAAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1313

```

```

Db      |||||
QY 703 TT-----CCTCTCCCATGTTACTCTCTTATGATTTTCTCTATCGTCAATATT 753
QY 1314 GGAATTTTCACTCATCTGTAATCCCATTTGCTATGATTTATGAATGAATACTTCAA 1373
Db      |||||
QY 754 GGAATTTTCACTCATCTGTAATCCCATTTGCTATGATTTATGAATGAATACTTCAA 813
QY 1374 AAAAATGTTTGTCTGCGAGTTTGTATTGATGATGATGATGATGATGATGATGATGAT 1433
Db      |||||
QY 814 AAAAATGTTTGTCTGCGAGTTTGTATTGATGATGATGATGATGATGATGATGATGAT 873
QY 1434 AGGCATGAAATTCAGGAATTTACATGATGCGGAAGAAAGCAAAAGTTTTCCTCAGAG 1493
Db      |||||
QY 874 AGGCATGAAATTCAGGAATTTACATGATGCGGAAGAAAGCAAAAGTTTTCCTCAGAG 933
QY 1494 AATCCAGTGGAGAAACC 1511
Db      |||||
QY 934 AATCCAGTGGAGAAACC 951

```

```

RESULT 15
AC093816 170227 bp DNA linear PRI 21-FEB-2002
LOCUS Homo sapiens BAC clone RP11-364P2 from 4, complete sequence.
DEFINITION AC093816 AC055821
AC093816.3 GI:18450197
VERSION HTG
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 170227)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 170227)
AUTHORS Shahid, S., Haakenson, W. and Creason, K.
TITLE The sequence of Homo sapiens BAC clone RP11-364P2
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 170227)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 170227)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (31-JAN-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 170227)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT On Jan 31, 2002 this sequence version replaced gi:15799650.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.edu
----- Summary Statistics
Center project name: H_NH0364P02
Drafting Center: WIBR
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RP11-11 human BAC library was made from the blood of one male donor, as described by Oseegawa, K., Woon, P.Y., Zhao, B., Frengen, E., Tatenno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-587P6. Actual start of this clone is at base position 73165 of RP11-587P6; actual end is at base position 170227 of RP11-364P2.

Data from AC093867 was used to finish this clone, AC093816.

Polymorphism has been identified between AC093867 and AC093816.

The sequence of AC055821 has been incorporated into AC093816.

FEATURES

source	Location/Qualifiers
repeat_region	1..230 /rpt_family="L1"
repeat_region	1644..1757 /rpt_family="L1"
repeat_region	1762..1946 /rpt_family="L1"
repeat_region	2511..2587 /rpt_family="L1"
repeat_region	2803..3200 /rpt_family="MaLR"
repeat_region	3216..3246 /rpt_family="AT_rich"
repeat_region	3249..6905 /rpt_family="L1"
repeat_region	6416..6459 /rpt_family="AT_rich"
repeat_region	6907..6927 /rpt_family="AT_rich"
repeat_region	6920..7561 /rpt_family="L1"
repeat_region	7562..7651 /rpt_family="MaLR"
repeat_region	7704..7899 /rpt_family="MaLR"
repeat_region	8069..8089 /rpt_family="AT_rich"
repeat_region	8260..8422 /rpt_family="MER1_type"
repeat_region	8905..9108 /rpt_family="MER1_type"
repeat_region	9164..9184 /rpt_family="AT_rich"
repeat_region	9484..9708 /rpt_family="MIR"
repeat_region	9903..10288 /rpt_family="ERV1"
repeat_region	10288..10471 /rpt_family="(TG)n"
repeat_region	10310..10396 /rpt_family="ERV1"
repeat_region	10474..10527 /rpt_family="ERV1"
repeat_region	11244..11372 /rpt_family="MER1_type"
repeat_region	11373..11515 /rpt_family="L2"
repeat_region	11581..11659 /rpt_family="L2"
repeat_region	12197..12356 /rpt_family="MER1_type"
repeat_region	13088..13442 /rpt_family="L1"
repeat_region	13706..14006 /rpt_family="Alu"
repeat_region	15009..15322 /rpt_family="Alu"
repeat_region	15294..15327 /rpt_family="(A)n"
repeat_region	15877..16112 /rpt_family="Alu"
repeat_region	16108..16144 /rpt_family="AT_rich"
repeat_region	16145..16309 /rpt_family="(TA)n"
repeat_region	16310..16344 /rpt_family="AT_rich"
repeat_region	16814..16916 /rpt_family="MIR"
repeat_region	16945..17421 /rpt_family="MaLR"
repeat_region	17427..17488 /rpt_family="MIR"
misc_feature	17507..17510 /notes="similar to Homo sapiens EST AA917407 (NID:G3057297)"
misc_feature	17698..17701 /notes="similar to Homo sapiens EST AA224463 (NID:G1845081)"
misc_feature	17878..18245 /notes="match to EST BE091143 (NID:G8481595)"
repeat_region	18407..18491 /rpt_family="MIR"
misc_feature	18456 /notes="similar to Homo sapiens EST AI078262 (NID:G3412670)"
misc_feature	18456 /notes="similar to Homo sapiens EST BE966526 (NID:G11772035)"
misc_feature	18536..19120 /notes="similar to Homo sapiens EST BE966526 (NID:G11772035)"
misc_feature	18536..18918 /notes="similar to Homo sapiens EST AI078262 (NID:G3412670)"
misc_feature	18537..19122 /notes="similar to Homo sapiens EST BG254743 (NID:G12764559)"
misc_feature	18537..19080 /notes="similar to Homo sapiens EST AW007850 (NID:G5856628)"
misc_feature	18537..18833 /notes="similar to Homo sapiens EST BI918363 (NID:G16182086)"

Misc_feature 18537.18548
/note="similar to Homo sapiens EST AW157456 (NID:g6228857)
au95g10.xl"

Query Match 28.7%; Score 693.4; DB 9; Length 170227;
Best Local Similarity 99.9%; Pred. No. 6.6e-131;
Matches 694; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGAGCGCCAGAGACCTAGCGTGGCGTCCAGCACCCAGACACCTGCGCGCGCTCG 60
DB 31654 GCCAGAGCGCCAGAGACCTCGCGTGGCGTCCAGCACCCAGACACCTGCGCGCGCTCG 31595

QY 61 CCTTAGGAGAGAGCAAGGAGAACTTATTGAAACCGCGAACTTTTGGTCACTGAG 120
DB 31594 CCTTAGGAGAGAGCAAGGAGAACTTATTGAAACCGCGAACTTTTGGTCACTGAG 31535

QY 121 ATCGAGTCTCCAGTCTTTGGCTTCCCGCTCTTTATCGTGGGTTTGTATCCCTGAGCTG 180
DB 31534 ATCGAGTCTCCAGTCTTTGGCTTCCCGCTCTTTATCGTGGGTTTGTATCCCTGAGCTG 31475

QY 181 CTCTCCTTTCCGAACTCCCGGGGTGAGCTTAGAGCCCTCCGCGCGCTGACTCCAG 240
DB 31474 CTCTCCTTTCCGAACTCCCGGGGTGAGCTTAGAGCCCTCCGCGCGCTGACTCCAG 31415

QY 241 AGTAGAGAGAGGAGCGGCTCCGCTGGTCCCGCAAGCCCTCGCTGCCCGCAGATG 300
DB 31414 AGTAGAGAGAGGAGCGGCTCCGCTGGTCCCGCAAGCCCTCGCTGCCCGCAGATG 31355

QY 301 CGGATGGCCAGCCAGTAGCGGCGGTGGCCCGGTCCTCCGGAGCGCACAGCAATGCAGG 360
DB 31354 CGGATGGCCAGCCAGTAGCGGCGGTGGCCCGGTCCTCCGGAGCGCACAGCAATGCAGG 31295

QY 361 CGCTTACATTAACCCCGAGAGTCTCTGGCTGCTGGGAGCCACCAACTGACGCGGG 420
DB 31294 CGCTTACATTAACCCCGAGAGTCTCTGGCTGCTGGGAGCCACCAACTGACGCGGG 31235

QY 421 AGCAGTTTCATCGTCTGTACCGGCTGGAGCCGCTCGTCTACCCCGAGAGCTGCCGGAC 480
DB 31234 AGCAGTTTCATCGTCTGTACCGGCTGGAGCCGCTCGTCTACCCCGAGAGCTGCCGGAC 31175

QY 481 GCGCCAAAGCTGGCCCTCGTCTACCGGCTGGTCTATCTTCCCGCTGGCGCTCTTTGGCA 540
DB 31174 GCGCCAAAGCTGGCCCTCGTCTACCGGCTGGTCTATCTTCCCGCTGGCGCTCTTTGGCA 31115

QY 541 ATGCTCTGGTGTCTACCGTGGTGAACCCGAGCAAGGCCATGCGCACCGTCAACCAATCT 600
DB 31114 ATGCTCTGGTGTCTACCGTGGTGAACCCGAGCAAGGCCATGCGCACCGTCAACCAATCT 31055

QY 601 TTATCTGCTCCTTGGCGCTCAGTACCTGCTCATCACTTCTTGTGCAFTCCCGTCAACA 660
DB 31054 TTATCTGCTCCTTGGCGCTCAGTACCTGCTCATCACTTCTTGTGCAFTCCCGTCAACA 30995

QY 661 TGCTCCAGAACATTTCCGACAACTGGCTGGGGGT 695
DB 30994 TGCTCCAGAACATTTCCGACAACTGGCTGGGGGT 30960

Search completed: October 29, 2003, 10:29:02
Job time : 8769 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 29, 2003, 07:12:19 ; Search time 4992 Seconds
(without alignments)
11757.875 Million cell updates/sec

Title: US-10-070-241B-2

Perfect score: 2415
Sequence: 1 gccagagcgccaggaccct.....aaaaaaaaaaaaaaagg 2415

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estc:*

9: gb_est1:*

10: gb_est2:*

11: gb_est3:*

12: gb_est4:*

13: gb_est5:*

14: gb_est6:*

15: em_estom:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rtd:*

26: em_gss_png:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	895.8	37.1	1642	11 AK051723	AK051723 Mus muscu
2	879.8	36.4	2864	11 AK048439	AK048439 Mus muscu
3	651.2	27.0	769	10 BG471265	BG471265 602512391
4	531.2	22.0	740	12 BT729969	BT729969 603350046

5	512.6	21.2	657	10	BB656182
6	507.2	21.0	730	10	BG169612
7	493	20.4	630	10	BB649191
8	478.2	19.8	635	10	BB663258
9	432.4	17.9	645	10	BB626475
10	414.6	17.2	611	13	BU462231
11	380.8	15.8	651	10	BB641290
12	376.6	15.6	726	13	BU453865
13	374.2	15.5	634	10	BB642180
14	370.8	15.4	866	13	BQ735771
15	324	13.4	888	13	BU458649
16	315.6	13.1	664	10	BB084541
17	309	12.8	432	9	AI307658
18	309	12.8	432	9	AI308124
19	305.2	12.6	665	9	AL878599
20	302.6	12.5	642	9	AL651383
21	285	11.8	546	28	AZ454676
22	252.6	10.5	609	9	AV760466
23	247.8	10.3	504	28	AQ215853
24	246.2	10.2	360	28	B91708
25	246.2	10.2	492	12	BM993628
26	246.2	10.2	669	29	AG155696
27	246	10.2	876	28	AQ746417
28	245.8	10.2	672	29	AG084721
29	244.2	10.1	540	13	BUG17996
30	244.2	10.1	921	28	AQ738768
31	243.8	10.1	465	9	AI613280
32	243.6	10.1	463	14	CA427664
33	243.6	10.1	633	13	EX471308
34	242.8	10.1	630	28	AQ528459
35	242.8	10.1	1050	14	CD523308
36	242.6	10.0	524	28	AQ278931
37	242.2	10.0	547	13	BU600474
38	241.6	10.0	664	29	AG040130
39	241.4	10.0	660	29	AG113120
40	241.4	10.0	687	29	AG113667
41	241.2	10.0	342	9	AW276435
42	241.2	10.0	434	28	AQ483420
43	241.2	10.0	924	28	AQ748907
44	241	10.0	471	28	AQ583262
45	240.8	10.0	381	10	BG272215

ALIGNMENTS

RESULT 1	AK051723	1642 bp	mrna	linear	HTC 05-DEC-2002
LOCUS	AK051723	Mus musculus 12 days embryo spinal ganglion cDNA, RIKEN full-length enriched library, clone:DI30067L02 product:hypothetical Rhodopsin-like GPCR superfamily containing protein, full insert sequence.			
DEFINITION	AK051723.1	GI:26342145			
ACCESSION	AK051723				
VERSION	AK051723.1	GI:26342145			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	1	Mammalia; Eutheria; Chordata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Itoh M., Konno H., Okazaki Y., Muramatsu M. and Hayashizaki Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)			
AUTHORS	Carninci P. and Hayashizaki Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE	2				
AUTHORS	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Konno H., Okazaki Y., Muramatsu M. and Hayashizaki Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				

FEATURES	source	Location/Qualifiers
11042159		1. 1642
3		/organism="Mus musculus"
REFERENCE		/mol_type="mRNA"
AUTHORS		/strain="G57BL/6J"
		/db_xref="FANTOM_DB.D130067L02"
		/db_xref="FANTOM.D130067L02"
		/clones="D130067L02"
		/tissue_type="spinal ganglion"
TITLE		/clone_lib="RIKEN full-length enriched mouse cDNA library"
		/dev_stage="12 days embryo"
88. 1185		
CDS		/note="unnamed protein product; hypothetical Rhodopsin-like GPCR superfamily containing protein (InterPro IPR000276, evidence: InterPro) putative"
		/codon_start=1
		/protein_id="BAC34735.1"
		/db_xref="GI:26342146"
		/translation="MLIFALPAGNALVVVTRSKAMRVTVTFICSLALSDLIIVF FCIPTVTLQNVSDTGLGAFICRMVQVCTAIETLTCTIAVERHQGLVHFFKWK ROYTNORAFMTLGVWLVVLIIGSPMMHVRLEIKYDFLYEKHEICCELEWSPVHOK IYTFILVTLFLPLLLSVLYGKIGVYELMKKIGDGLVLRTHGEMFKPKKR AVIMVTVVWLFAVCAPFHVHMIEVSNFEKYEYDEVTKMTFAIVOIGFENSIGN PLIYALMNEFKNFVSNVNCVCIUKETPSSARKHGSSGANVHRRALAAENPVEIK GEATGGSNIDIKWCEQEKKRRSKVASCLP"
BASE COUNT	466 a	359 c 372 g 445 t
ORIGIN		
	Query Match	37.1%; Score 895.8; DB 11; Length 1642;
	Best Local Similarity	82.5%; Pred. No. 9.5e-13;
	Matches 1091; Conservative	0; Mismatches 217; Indels 15; Gaps 5;
QY	424	AGTTGATGCTCTGTACCGGCTGGACCGCTGCTCTACACCCAGAGCTGCCGGAGCGG 483
DB	2	AGTTGATGCTCTGTACCGGCTGGACCGCTGCTCTACACCCAGAGCTGCCGGAGCGG 61
QY	484	CCAAGCTGGCCCTCGTGTCTACCGGCTGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 543
DB	62	CCAGGCTGGCCCTCGTGTCTGCTGGCATGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 121
QY	544	CTCTGGTGTCTGTACCGGCTGGACCGCTGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 603
DB	122	CCCTGGTGTCTGTACCGGCTGGACCGCTGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 181
QY	604	CTCTGGTGTCTGTACCGGCTGGACCGCTGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 663
DB	182	CTCTGGTGTCTGTACCGGCTGGACCGCTGCTCATCTTGGCCCTGGCGCTCTTTGGCAATG 241
QY	664	TCAGAACATTTCCGACAACTGGCTGGGGGGTGTCTTCATTTGCAAGATGTGCGAATTTG 723
DB	242	TCAGAACATTTCCGACAACTGGCTGGGGGGTGTCTTCATTTGCAAGATGTGCGAATTTG 301
QY	724	TCAGTGTACCGCTGTGTGACAGAAATCCTCCTATGACCTGCTGCTGTGGAAGGC 783
DB	302	TCAGTGTACCGCTGTGTGACAGAAATCCTCCTATGACCTGCTGCTGTGGAAGGC 361
QY	784	ACCAGGAGCTGTGCACTCTTTTAAATGAGTGGCAATACCAACCGAAGGCTTTCA 843
DB	362	ACCAGGAGCTGTGCACTCTTTTAAATGAGTGGCAATACCAACCGAAGGCTTTCA 421
QY	844	CAATGCTAGGTGTGTCTGGCTGGGGAGTCACTGATGATCACCATGTGGGACGTGC 903
DB	422	CAATGCTAGGTGTGTCTGGCTGGGGAGTCACTGATGATCACCATGTGGGACGTGC 481
QY	904	ACCACTTGAGATCAATATGACTTCCTATATGAAAGGAACACATCTGCTGCTTAGAAG 963
DB	482	ACCACTTGAGATCAATATGACTTCCTATATGAAAGGAACACATCTGCTGCTTAGAAG 541
QY	964	AGTGGACGAGCCCTGTGACCAAGAGATCTACACCACTTTCATCTTGTCTCTCTTCC 1023
DB	542	AGTGGACGAGCCCTGTGACCAAGAGATCTACACCACTTTCATCTTGTCTCTCTTCC 601

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuyama, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.		
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer		
Genome Res. 10 (11), 1757-1771 (2000)		
20530913		
11076861		
4		
Kawai, J., Shinagawa, A., Shibata, K., Yoshino, M., Itoh, M., Ishii, Y., Arakawa, T., Hara, A., Fukunishi, Y., Konno, H., Adachi, J., Fukuda, S., Aizawa, K., Izawa, M., Nishi, K., Kiyosawa, H., Kondo, S., Yamanaka, I., Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischnann, W., Gaasterland, T., Gissi, C., King, B., Kochiwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L.M., Stauber, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Anono, H., Baldarelli, R., Bareh, G., Blake, J., Boffelli, D., Bojunga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hume, D.A., Kaniyama, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarelli, J., Mombaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seva, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whittaker, C., Wilming, L., Wyshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohtsuki, S., and Hayashizaki, Y.		
Functional annotation of a full-length mouse cDNA collection		
Nature 409 (6821), 685-690 (2001)		
21085660		
11217851		
5		
The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.		
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs		
Nature 420, 563-573 (2002)		
6 (bases 1 to 1642)		
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Kikuchi, T., Kishi, P., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Soabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akai, S., Takeda, Y., Tanaka, T., Tomaru, A., Toyota, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.		
Direct Submission		
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)		
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.		
Please visit our web site for further details.		
URL: http://genome.gsc.riken.go.jp/		
URL: http://fantom.gsc.riken.go.jp/		

Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.go.jp/>
URL: <http://fantom.gsc.riken.go.jp/>

FEATURES

source

Location/Qualifiers
1. 2864
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="ANTOM DB.C130060K24"
/db_xref="taxon:10090"
/clones="Cl30060K24"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
64. 1161
/note="unnamed protein product; hypothetical
Rhodopsin-like GPCR superfamily containing protein
(InterPro:IPR000276, evidence: InterPro)
putative"
/codon_start=1
/protein_id="BAC33337.1"
/db_xref="GI:26339332"
/translation="MLIPALALPGLNLLVYVYVTRSKMRTVYNIFICSLALSDLIVF
FCIPVLMQNVSLGGGAFICXMPVFCVTAIVELINTCIABVHGLVHPFK
ROYNORAFMLGVMLVAILIGSPMVHQRLEIKYDFEYKEHICCLBEWSPVHOK
IYTFILVFLFLVLLSLYKIGYELKIKRIGDGLVIRKIFKFKARKKR
AVIMVTVVLFAVCWAPFHVHMYIENYFEYDEVTKMIFAIQIIGFNSICN
PIYALMNEFKNPFVAVCYCIKVPSSARKHGSGAMVHRRRAKLAARENPEIK
GEAFGGSNIDIKWCEQPEKKRKRKSCFL"
BASE COUNT 839 a 609 c 602 g 814 t

CDS

Query Match 36.4%; Score 879.8; DB 11; Length 2864;
Best Local Similarity 82.5%; Pred. No. 1.4e-12;
Matches 1072; Conservative 0; Mismatches 212; Indels 15; Gaps 5;
QY 448 GACCGCTCGTACACCCACGAGCTGCGGAGCGCCAGAGCTGGCCCTCGTGTACCG 507
DB 2 GGCACCTGGTCTCACCCCGCAGCTTCCGCGCGCCGAGCTGGCCCTCGTGTGCG 61
QY 508 GCGTGTCTATCTCGCCCTGCGCTCTTTGGCAATGCTCTGGTGTCTACGTGTGACCC 567
DB 62 GCATGCTCATCTTTGGCCCTGGCGCTCTTCGCAACGCCCTGGTAGTCTATGTGTGACC 121
QY 568 GCAGAGGCGATGGGACCGCTACCAACATCTTTATCTGCTCTGCGGCGCTCAGTGACC 627
DB 122 GCAGAGGCGATGGGACCGCTACCAACATCTTTATCTGCTCTGCGGCGCTCAGTGACC 181
QY 628 TGCTCATCACCTTCTCTGCAATTCGCGTACCATGCTCCAGAACATTTCCGACAACTGGC 687
DB 182 TGCTCATCGTCTTCTCTGCAATTCGCGTACCATGCTCCAGAACATTTCCGACAACTGGC 241
QY 688 TGGGGGTGCTTCAATTTGCAAGTGTGTCATTTGTCAGTCTACCGCTGTGTGACAG 747
DB 242 TGGGGGTGCTTCAATTTGCAAGTGTGTCATTTGTCAGTCTACCGCTGTGTGACAG 301
QY 748 AAATCCTCAGTATGACCTGCAATTTGCTGGAAGGACCCAGGAGCTTGTGCATCCTTTA 807
DB 302 AAATCCTTACTATGACCTGCAATTTGCTGGAAGGACCCAGGAGCTTGTGCATCCTTTA 361
QY 808 AAATGAAGTGGCAATACACCAACGAGGCTTTTCAATGCTAGGTGTGTGCTGGCTGG 867
DB 362 AAATGAAGCGGAGTACACCAATCAAGAGCTTTTCAATGCTAGGTGTGTGCTGGCTGG 421
QY 868 TGGGAGTATCTAGGATCACCATGCTGGCAGCTGTCACACATCTGAGATCAATGACT 927
DB 422 TGGGAGTATCTAGGATCACCATGCTGGCAGCTGTCACACATCTGAGATCAATGACT 481
QY 928 TCCTATATGAAGAAACACATCTGCTGTAGAGAGTGGACCGCCCTGTGCACGAGA 987

DB 482 TCCTATATGAAGAAACACATCTGCTGCTGGAAGTGGAGCAGCCCGTGCACAGA 541
QY 988 AGATCTACACCACTTCATCTTGTCTCATCTCTTCTCTCTCTCTCTCTCTCTCTCT 1047
DB 542 AGATCTACACCACTTCATCTTGTCTCATCTCTTCTCTCTCTCTCTCTCTCTCTCT 601
QY 1048 TTCTGTACAGTAAATTTGGTATGAACTTTGGATTAAGAAAGAGTTGGGATGTTTCAG 1107
DB 602 TCCTCTACGGGAAATCGGTTATGAGCTTTGGATCAAGAAAGAAATCGGGGATGGCTCAG 661
QY 1108 TGCTTCGAACTATTCATGGAAGAAATGTCGCAAAATAGCAGAGAAAGAAACGAGCTG 1167
DB 662 TGCTCCGAACTATTCATGGAAGAAATGTCGCAAAATAGCAGAGAAAGAAACGAGCTG 721
QY 1168 TCATTTATGCTGACAGTGGTGGCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1227
DB 722 TGATCATGCTGGTGGAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 781
QY 1228 TCCATATGATGATTAATACAGTAATTTTGAAGAAAGAAATGATGATGTCACAAATCAAGA 1287
DB 782 TTCATATGATGATTAATACAGTAATTTTGAAGAAAGAAATGATGATGTCACAAATCAAGA 841
QY 1288 TGATTTTGTCTATCGTCAAAATTTATGGAATTTTCCAACTCCATCTGTAATCCCATGCT 1347
DB 842 TGAATTTGTCTATAGTCAAAATTTTGAATTTTCAACTCCATCTGTAATCCCATGCT 901
QY 1348 ATGCAATTTGATGAAGAACTTCAAAAGAAATTTTGTCTGTCAGATTTGTTATGTCATAG 1407
DB 902 ATGCAATTTGATGAAGAACTTCAAAAGAAATTTTGTCTGTCAGATTTGTTATGTCATAG 961
QY 1408 TAAATAAAACCTTCTCTCCAGCAAAAGGATGAAATTTGAGAAATTTTCAATGATGCGGA 1467
DB 962 TGAAGGAAACCTTCTCTCCAGCAAAAGGATGAAATTTGAGAAATTTTCAATGATGCGGA 1021
QY 1468 AGAAGCAAGCTTTTCCCTCAGAGAGATCCAGTGGAGGAAACCAAGAGGAGGATTCAC 1527
DB 1022 GAGGGCAAGCTTTAGCTGCAAGAGAGATCTCTGTAGAG---ATCAAAGGAGAGCAATTTG 1078
QY 1528 GTGATGGCAACATTAAGTCAAAATTTGTGTGAACAGACAGAGGAGGAGAA---AAAGCTC 1583
DB 1079 GGGGCAAGCAATCGATATCAAGTGTGTGAACAGGAGGAGGAGGAGGATCAAA 1138
QY 1584 AAACAGACATCTTCTCTTTAGGTGTGAAGTGGCTGAGATTTCTCTTTAGACAGTGGG 1643
DB 1139 AAGTGGCACTTGTCTCTTTAGTTCGAA---TTTCTGAGAGCTCTGTGTAGAC---G 1192
QY 1644 CATTAAATTAACAAATATCTTCAATTAATTAATGCCCCTTCAGATTTGTAACCAAGAGAAA 1703
DB 1193 TGAACACTGTACCAATGCTTTCAGATGAGTATCTGTCTACTGTATGTAATCGAA---AGAAA 1250
QY 1704 TTAATTTGAGCAAGGTCAAATATCTCTTTTAATCTTAA 1742
DB 1251 TGAATTTGAGCAAGGTCAGAGAGCTTTTCAATTAATAAAA 1289

RESULT 3

BG471265

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG471265 769 bp mRNA linear EST 21-MAR-2001
602512391F1 NIH_MGC_16 Homo sapiens cDNA clone IMAGE:4634903 5',
mRNA sequence.
BG471265
BG471265.1 GI:13403540
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 769)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCMI391 row: f column: 24
High quality sequence stop: 765.
Location/Qualifiers
1. .769
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4634903"
/tissue_type="retinoblastoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 18"
/note="Organ: eye; Vector: pOTB7; Site: 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
Note: this is a NIH_MGC library."

BASE COUNT 238 a 177 c 165 g 189 t
ORIGIN

Query Match 27.0%; Score 651.2; DB 10; Length 769;
Best Local Similarity 96.4%; Pred. No. 4e-07;
Matches 740; Conservative 0; Mismatches 20; Indels 8; Gaps 7;

QY 1564 CAGAGGAGAGAAAAGCTCAACGACATCTTGCTCTCTTTAGGCTGAACTGGCTGAGA 1623
DB 2 CAGAGGAGAGAAAAGCTCAACGACATCTTGCTCTCTTTAGGCTGAACTGGCTGAGA 61

QY 1624 ATTCTCTTTAGACAGTGGGCATTAATTAACAATATCTTCAATATTAATGCTTCAG 1683
DB 62 ATTCTCTTTAGACAGTGGGCATTAATTAACAATATCTTCAATATTAATGCTTCAG 121

QY 1684 ATTGTAAACCAAGAGAA-AATTATTTCAGCAAGGTCAAATACTCTTTTAATCTTAA 1742
DB 122 ATTGTAAACCAAGAGAAATATTTCAGCAAGGTCAAATACTCTTTTAATCTTAA 181

QY 1743 GATGATGACAAGAGAAAACAATCATGTTTCCATTAAAAA-AATGACAGAGCTAGTCC 1801
DB 182 GATGATGACAAGAGAAAACAATCATGTTTCCATTAAAAACAATGACAGAGCTAGTCC 241

QY 1802 AAGTCAGTGTATTTACACCAATNGATCAACAATCATTTACANATTTCTGTTCCTT 1861
DB 242 AAGTCAGTGTATTTACACCAATNGATCAACAATCATTTACANATTTCTGTTCCTT 301

QY 1862 CTCATTCCCACTGCTTCATCTGACTAGCTTAAAAAGCAACATGAAGCCAGGCAAGG 1921
DB 302 CTCATTCCCACTGCTTCATCTGACTAGCTTAAAAAGCAACATGAAGCCAGGCAAGG 361

QY 1922 TGGCTCATGCTGTATATCCAGACATTTGGGAGGCTTAGACGGGGGATCAGAGGTCTAG 1981
DB 362 TGGCTCATGCTGTATATCCAGACATTTGGGAGGCTTAGACGGGGGATCAGAGGTCTAG 421

QY 1982 GAGATCAAAACCATCTGTGCTAACCGGTGAACCCCATCTCTGTCTAAATAATACAAAAT 2041
DB 422 GAGATCAAAACCATCTGTGCTAACCGGTGAACCCCATCTCTGTCTAAATAATACAAAAT 481

QY 2042 TAGCCGGGGCTGTGTGGCGGACCTCTAGTCCAGCTACTTGGAGCCCTCAGCGGGAG- 2100
DB 482 TAGCCGGGGCTGTGTGGCGGACCTGTAGTCCAGCTACTTGGAGCCCTCAGCGGGAG 541

QY 2101 AATGTTGTGAACCCGGAGGCGGAGCTTGCAGTGTATCCGAGATCATGCGCACTCCCA 2160
DB 542 AATGTTGTGAACCCGGAGGCGGAGCTTGCAGTGTATCCGAGATCATGCGCACTCCCA 601

QY 2161 -GCCTGGGCGAAGACGAGACATCCCGCTCTCAAAAAAATTTTTTGAAAAATTCGTAA 2219

Db 602 GGCCTGGCGAAGAGCGAGACTCCCGCTCTCAAAAAAATATATTGGAATAATTCGTAA 661
QY 2220 ACC--ATACTTTAGATATTTAGTGTATTTTAAAAA-ATCTTGTACAGAAATCAGGG 2276
Db 662 ACCCACTATTTTACAGATTTCTTCAGTGTATTTTAAAAACATCTCGTACAGAAATCAGGG 721
QY 2277 TTCTTAG-CTAGCAGTTTTTCTCCACGCGAGTCACTGTAATGTGACTA 2323
Db 722 TCTTAGCTAGCAGATTAATCTCCACGCGAGTCACTGTAATGTGACCA 769

RESULT 4
LOCUS B1729969
DEFINITION 60350046F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5357699 5', mRNA sequence.
ACCESSION B1729969
VERSION B1729969.1 GI:15706969
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 740)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgpbbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLCMI391 row: k column: 12
High quality sequence stop: 712.
Location/Qualifiers
1. .740
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:5357699"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_94"
/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies."
Note: this is a NIH_MGC library."

BASE COUNT 150 a 225 c 184 g 181 t
ORIGIN

Query Match 22.0%; Score 531.2; DB 12; Length 740;
Best Local Similarity 84.7%; Pred. No. 0.00018;
Matches 596; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

QY 354 ATGAGGGCGCTTAACATTAACCCCGAGAGAGTCTCTCGGCTGCTGCGGAGCACAACTG 413
DB 37 ATGAGGGCGCTCAACATCAACCGGAGCAGTTTTCCGGCTGCTGAGCGGCACAACTG 96

QY 414 ACGGGGAGAGCTTCATCGCTCTGTACCGGCTGGACCGCTCGTCTACACCCGAGCTG 473
DB 97 ACTCGGGAACAGTTCATTCATCGCTATGGGCTGGACCGCTGCTCTACACTCCGAGCTG 156

QY 474 CCGGAGCGCGCAAGCTCGGCCCTCGTGTCTACCCGCGGTCTCATCTTCGCGCTCGCGCTC 533
DB 157 CCGGCGCGCGCTAACTGGCCCTTTCGCTGGCTGAGCAGCTCATTTTTCGCGCTCGCGCTC 216

QY 534 TTTGGCAATGCTCTGTGTGTTTCTAGTGTGTGACCGCGAGGCCATCGGCACCGTCAAC 593

```

Db      217 TTTGGCAACTCTCTGGTCACTATGTTGGTACCGCGCAAGGCCATGCGCAGCTCAACC 276
QY      |||||
594 AACATCTTTTCTCTCTCTGGGCTCAGTGAACCTGCTCATCACTTCTTCTGCACTTCC 653
Db      |||||
277 AACATCTTCTCTCTCTCTGGCACTCAGTGAATCTGCTCATTTGCTTCTTCTGCACTCCC 336
QY      |||||
654 GTACACATGCTCAGAACATTTTCGACAACTGGCTGGGGGTGCTTTCAATTTCAAGATG 713
Db      |||||
337 GTACAGATGCTCCAGAACATCTCCGACAGTGGCTGGTGGTGGCTTCACTTCAAGATG 396
QY      |||||
714 GTGCCATTTCTCCAGTCTACCGCTGTGTGTCAGACAAATCTCACTATGACCTGATGCT 773
Db      |||||
397 GTGCCCTTGTGTCAGTCCATCTGCTGTGTGACGGAAATCTCCACCATGACTTTCATCGT 456
QY      |||||
774 GTGGAAGGCACACGAGGACTTGTGCATCTCTTTAAATGAAGTGGCAATACACCAACCGA 833
Db      |||||
457 GTTGAGAGGCACCAAGGACTCATCCATCTCTTTAAATGAAGTGGCAATACACCAACCGA 516
QY      |||||
834 AGGGCTTTCACATGCTAGTGTGGTCTGGCTGGTGGCAGTCACTAGGATCAACCATG 893
Db      |||||
517 AGGGCTTTCACATCTTGGGTGGTCTGGTGGTGGCAGCCTATCATCTAGGATCAACCATG 576
QY      |||||
894 TGGCAGTGTCAACAACTTGAGATCAAAATATGACTTCTATATGAAAGGACACATCTGC 953
Db      |||||
577 TGGCAGTGTCAACAACTTGAGATCAAAATATGACTTCTATATGAAAGGACACATCTGC 636
QY      |||||
954 TGCTTAGAGAGTGGACCGCCCTGTGCACGAGAGATCTACACCACTTCACTCTCTGTC 1013
Db      |||||
637 TGTATGGAAGTGGGCCAGACCCATGACAGAGATCTACACCACTTCACTCTCTGTC 696
QY      |||||
1014 ATCCTCTTCTCTCTCTCTTATGATGATGCTTATTTCTGTACAG 1057
Db      |||||
697 ATCCTCTTCTCTCTCTCTTATGATGATGCTTATTTCTGTACAG 740

RESULT 5
LOCUS   BB656182
DEFINITION BB656182 RIKEN full-length enriched, 12 days embryo spinal ganglion
          BB656182
          BB656182.1 GI:16490010
KEYWORDS EST.
SOURCE  Mus musculus (house mouse)
ORGANISM
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
          1 (bases 1 to 657)
          Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
          Hiramoto, K., Horii, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda,
          M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M.,
          Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki,
          D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H.,
          Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T.,
          Muramatsu, M. and Hayashizaki, Y.
          RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
TITLE   Unpublished
JOURNAL
COMMENT Contact: Yoshihide Hayashizaki
          Laboratory for Genome Exploration Research Group, RIKEN Genomic
          Sciences Center (GSC), Yokohama Institute
          The Institute of Physical and Chemical Research (RIKEN)
          1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          Tel: 81-45-503-9222
          Fax: 81-45-503-9216
          Email: genome-res@gsr.riken.go.jp,
          URL: http://genome.gsc.riken.go.jp/
          Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh,
          M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
          Normalization and subtraction of cap-trapper-selected cDNAs to
          prepare full-length cDNA libraries for rapid discovery of new
          genes. Genome Res. 10 (10), 1617-1630 (2000)
          wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,

```

Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
 Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Location/Qualifiers
 1. 657
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="D130067L02"
 /tissue_type="spinal ganglion"
 /dev_stage="12 days embryo"
 /lab_host="DH10B"
 /clone_lib="RIKEN full-length enriched, 12 days embryo spinal ganglion"
 /note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5', GAGAGAGAGCGCGCACTCGACTGATTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTTCTCGAGTTAATTAATTAATCCCGCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."
 BASE COUNT 133 a 201 c 159 g 164 t

Query Match 21.2%; Score 512.6; DB 10; Length 657;
 Best Local Similarity 86.4%; Pred. No. 0.00049;
 Matches 566; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
 QY 424 AGTTTCATCGCTCTCTACGGTGGCGTGGCTGCTTACACCCAGAGTGTCCGGACGCG 483
 Db 2 AGTTTCATCGCGCACTATGTTGGTGGCGCACTGGTGTCTCACCCCGCAGCTTCCCGGCGC 61
 QY 484 CCAAGCTGGCCCTCGTGTCTACCGCGCTGCTCATCTTCGCCCTGCGCTCTTTGSCATG 543
 Db 62 CCAGCTGGCCCTCTCTGCTGTGGCAATGCTCATCTTTGCTGGCGCTCTTCGGCAACG 121
 QY 544 CTCGTGGTGTCTACGTGGTGACCCGCGCAGGCGCATGCGCAGCGTCAACCAATCTTTA 603
 Db 122 CCCCTGGTAGTCTATGTGTTGACCCGCGCAGGCGCATGCGCAGCGTCAACCAATCTTCA 181
 QY 604 TCTGCTCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGATTCCTCGGTCAACATGC 663
 Db 182 TCTGCTCTTGGCGCTCAGTGACCTGCTCATCACCTTCTTCTGATTCCTCGGTCAACATGC 241
 QY 664 TCCAGAACATTTCCGACAACTGGCTGGGGGTGCTTTTCATTTCAAGATGTTGCCATTG 723
 Db 242 TCCAGAACATGCTCGACACCTGGCTGGGGGTGCTTTTCATTTCAAGATGTTGCCATTG 301
 QY 724 TCCAGTCTACCGCTGTTTGTGACAGAAATCCTCACTATGACCTGCTGTTGGAAAGGC 783

Query Match 20.4%; Score 493; DB 10; Length 630;
Best Local Similarity 86.5%; Pred. No. 0.0014;

1

1

1

, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

BASE COUNT	144 a	153 c	134 g	180 t	
ORIGIN					
Query Match	17.2%;	Score 414.6;	DB 13;	Length 611;	
Best Local Similarity	81.1%;	Pred. No. 0.074;			
Matches 494;	Conservative	0;	Mismatches 114;	Indels 1;	Gaps 1;
QY	518	CTTGGCGCTGGCGCTCTTTGGCAATGCTCTGGTGTCTTACGTGGTGAACCGCAGCAAGGC	577		
Db	3	CTTGGCGCTCACGCTTTTCGGCAACTCCCTGGTGTCTTACGTGGTGAACCGCAGC-AGGC	61		
QY	578	CATGGCGACCGTCACCAACAATCTTTATCTGCTCCTCTGGCGCTCAGTGACACTGCTGCATCAC	637		
Db	62	CATGAGGACCGTCACCAACAATCTTCATCTGCTCCTGGCGCTCAGCGACTCTCTCATCGC	121		
QY	638	CTTCTTCTGCAATCCCGTCCACATGCTCCAGAAATTTCCGACAACATGGCTGGGGGGTGC	697		
Db	122	CTTCTTCTGGTCCCTCTCCACATGCTGCAGAACATCTCCTCCGAGTGGCTGGCGGGTGC	181		
QY	698	TTTCANTTGAAGATGGTGCCATTTGTCAGTCTACCGCTGTGTGACAGAAATCTCTCAC	757		
Db	182	CTTCGGCTTGAAGATGGTACCATTTGTTCAATCCACTGCCATTGTAATCAGAGATCTCTTAC	241		
QY	758	TATGACCTGCATTCGCTGTGGAAGAGCACAGGGAATTTGTGCATCCTTTTAAATGAAGATG	817		
Db	242	AATGACCTGCATCGCTGTGGAAGAGCACAGGGAATTTGTGCATCCACTAAATATGAAGTG	301		
QY	818	GCAATACACCAACGGAAGGCTTTCACAATGCTAGGTGTGGTCTGGGCTGGTGGCAGTCAAT	877		
Db	302	GCAGTACACCAATAAAGGGCTTTACGATGCTTTGGCATATGCTGTGTTGCTGGCCATTAT	361		
QY	878	CGTAGGATACCCATGTGGCAGCTGCAACAACATTTGAGATCAAAATATCACTTCCATATATGA	937		
Db	362	TGTTGGTCACTATGTGGCATGTGCACCGCTTGAGTTAAATATGACTTCTGTATATGA	421		
QY	938	AAAGGAACACATCTGCTGTTAGAAGAGTGGACCGAGCCCTGTGCACCAAGAGATCTACAC	997		
Db	422	AAAAGTGACATCTGTTGCTTGGGAAGATGGGCCAGTCCAACTTTATCAGAAGATATATAC	481		
QY	998	CACCTTCATCCTTGTCAATCCTCTTCCTCCCTTATGGTGAATGCTTTATTCGTACAG	1057		
Db	482	CACCTTTATACCTGTTATATCTCTCTTTTCCACTGATATTGATGCTTTTCTTATACAC	541		
QY	1058	TAAAAATGGTTATGAACCTTTTGGATATAAGAAAAGAGTTGGGGATGGTTCAGTGCTTCGAAC	1117		
Db	542	TAAAAATGGTTATGAACCTTGGATTAAGAAACAGATGGGAGATGCTTCAGTCTTCAAAC	601		
QY	1118	TATTCATGG	1126		
Db	602	CATTTCATGG	610		

RESULT 11
BB641290
LOCUS
DEFINITION
BB641290 651 bp mRNA linear EST 26-OCT-2001
BB641290 RIXEN full-length enriched, 10 days neonate cortex Mus
musculus cDNA clone A830015G12 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

, Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Ozakaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.

Riken Mouse ESTs (Arakawa, T., et al. 2001)

Unpublished

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Sushiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: <http://genome.gsc.riken.go.jp/>

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Iishikawa, R., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the School in the Community	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of Socioeconomic Status on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Impact of Technology on Education	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Importance of Parental Involvement	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Effect of Teacher Expectations on Student Achievement	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp/
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh
M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subfraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
Wagai, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
Watanuki, M., Yoneda, Y., Iisuiawa, T., Ozawa, K., Tanaka, T., Matsura,
S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y., and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome* 12, 673-677 (2001)
Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
SOURCE

```

FEATURES
source
Location/Qualifiers
1. 651
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A830015G12"
/tissue_type="cortex"
/dev_stage="10 days neonate"
/lab_host="DH10B"
/clone_lib="RIKEN full-length enriched, 10 days neonate
cortex"
/note="Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGAGAGAGCCAGAGAGCTTTTTTTTTTTTNN 3'], cDNA was
prepared by using trihalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5', GAGAGAGAGATTCGAGTTAATTAATATCCCCCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
PhC I."
131 a 193 c 151 g 172 t 4 others
BASE COUNT
ORIGIN

```

BASE COUNT		FDC I.	
ORIGIN		a	c
131	a	193	c
151	g		
172	t		
4	others		

Query Match	15.8%; Score 380.8; DB 10; Length 651;
Best Local Similarity	84.7%; Pred. No. 0.39;
Matches	Conservative 0; Mismatches 77; Indels 0; Gaps 0;
Qy	413 GACGGGGAGAGATTTCATCGCTCTGTACCGCTGGGACCGCTGCTCTACACCCAGAGCT 472
Db	
Qy	3 GACTCGGGAAAGATTTCATTCGCTATGGCTGGACCGCTGGTCTACACTCCGGAGCT 62
Db	
Qy	473 GCCGGGACGCCCAAGCTGGCGCTCTGTCTCACCGGCTGCTCATCTTTGGCCCTGGCGCT 532
Db	
Qy	63 GCCCGCGCGCTAACTGGGCTTTGGCTGGAGCACTCAATTTTGGCCCTGGCGCT 122
Db	
Qy	533 CTTTGGCAATGCTCTGGTGTTCTACGTGGTGACCCGACGAGCCCATGGCACCGTCAC 592
Db	
Qy	123 CTTTGGCAACTCTCTGGTCACTATGTGGTGAACCGCAGCAAGGCCATGGCACCGTCAC 182
Db	
Qy	593 CAACATCTTTATCTGCTCTTGGGCTCAGTGACCTGCTCATCACTTTCTTCTGCATTC 652
Db	
Qy	183 CAACATCTTTCATGCTCTCTGGCACTCAGTGATCTGCTATTGCTCTTCTTCTGCATCCC 242
Db	
Qy	653 CGTCACCATGCTCCAGAAACATTTCCGACCACTGGCTGGGGGGTCTTTTCATTTCGAAGAT 712
Db	
Qy	243 CGTCACGATGCTCAGAAACATCTCCGACAGTGCTGGTGGTGCTTCATCTCAAGAT 302
Db	
Qy	713 GGTGCCATTTGTCAGTCTACCGCTGTTGTGACAGAAATCCTCACTATGACCTCATATGC 772
Db	
Qy	303 GGTGCCCTTCGTCCAGTCCACTGCTGTTGTGACGAAATCCTCACTATGACCTCATATGC 362
Db	
Qy	773 TGTGGAAGGCACGAGGACTTGTGCATCCTTTTAAATGAAGTGGCAATACACCAACCG 832
Db	
Qy	363 TGTGTGAGAGGCACCAAGGACTCATCCATCCTTTTAAATGAAGTGGCAGTACACTACCCG 422
Db	
Qy	833 AAGGGCTTTTCAATAGCTAGTGTTGGTCTGGCTGGTGGCAGTCACTCGTAGATCACCCAT 892
Db	
Qy	423 AAGGGCTTTTCAATCTTGGTGGTGGTCTGGTGGCAGCCATCATCTAGATCACCCAT 482
Db	
Qy	893 GTGCAAGTGCACAACTTGAGAT 916
Db	
Qy	483 GTGCAAGTGCACAACTTGAGAT 506
Db	
RESULT 12	
BU453865	
LOCUS	603770134F1 CSEQREN14 Gallus gallus mRNA linear EST 29-NOV-2002
DEFINITION	sequence.
ACCESSION	BU453865
VERSION	BU453865.1 GI:25943176
KEYWORDS	EST.
SOURCE	Gallus gallus (chicken)
ORGANISM	Gallus gallus
REFERENCE	Eukaryotes; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
TITLE	Phasianinae; Gallus.
JOURNAL	1 (bases 1 to 726)
MEDLINE	Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
PUBMED	Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
COMMENT	A Comprehensive Collection of Chicken CDNAs
	Curr. Biol. 12 (22), 1965-1969 (2002)
	223335534
	12445392
	Contact: Simon Hubbard
	Department of Biomolecular Sciences
	University of Manchester Institute of Science and Technology (UMIST
)
	PO Box 88, Manchester, M60 1QD, UK
	Tel: 01612008930
	Fax: 01612360409
	Email: Simon.Hubbard@umist.ac.uk.
FEATURES	Location/Qualifiers
source	1. 726
	/organism="Gallus gallus"
	/moltype="mRNA"

```

Db      661 CACTGACTCGGATGAGCCGAGGAGGCGCTCAGTGATGCGCAACATTGAGTGAAGTTC 720
      1554 TGTGA 1558
      721 TGTGA 725

RESULT 13
LOCUS   BB642180                               634 bp mRNA linear EST 31-AUG-2001
DEFINITION
clone A93003A06 5', mRNA sequence.
ACCESSION
BB642180
VERSION
BB642180.1 GI:15401798
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 634)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda
M., Kova,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okada
Okazaki,Y., Okido,T., Saico,R., Sakai,C., Sakai,K., Sano,H., Sasaki
D., Shibata,K., Shingawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T., et al. 2001)
Unpublished
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh
M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new
genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi,K., Fujiwaki,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,
Watahiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and
Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer. Genome Res.
10 (11), 1757-1771 (2000)
Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara
Y. and Hayashizaki,Y.
Computer-based methods for the mouse full-length cDNA
encyclopedia: real-time sequence clustering for construction of a
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Yamanaka,I., Kiyosawa,H., Kondo,S., Saito,T., Shinagawa,A., Aizawa
K., Fukuda,Y., Hara,A., Itoh,M., Kawai,J., Shibata,K., Arakawa,T.,
Ishii,Y. and Hayashizaki,Y.
Mapping of 19032 mouse cDNAs on mouse chromosomes. J. Struct.
Func. Genomics 2 pre, L72-L86 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp) for
further details.
e mouse tissues.
FEATURES
Location/Qualifiers
1..634
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A93003A06"
/tissue_type="retina"
/dev_stage="adult"
/lab_host="DH10B"

```

```

/clone lib="RIKEN full-length enriched, adult retina"
/notes Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'
GAGAGAGAGAGATCCAGAGACTCTTTTCTTTTCTTTT 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 20.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence [5'GAGAGAGAGATTCGAGTTAATAATTAATATATCCCCCCCC
3']. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pBluescript KS(+) after bulk excision from Lambda
FLC I. -Retina RNA was provided by Stefano Gustinich,
Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA02115, USA, whose assistance we
gratefully acknowledge."
BASE COUNT      131 a      189 c      167 g      147 t
ORIGIN
Query Match      15.5%; Score 374.2; DB 10; Length 634;
Best Local Similarity 80.7%; Pred. No. 0.55;
Matches 474; Conservative 0; Mismatches 108; Indels 5; Gaps 3;
QY      354 ATCGAGCGCTTAACATTACCCGGAGCAGTTCTCGGCTGCTGGGGACCAACCTG 413
      50 ATCGAGCGCTCAACATCACCGGAGCAGTTTCCCGGCTGCTGAGCGGCAACCTG 109
      414 ACCCGGAGCAGTTTCATCGCTCTGTACCGGCTCGGACCGCTCGTCTACACCCAGAGCTG 473
      110 ACTCGGAGCAGTTTCATCGCTATGCGCTGCGACCGCTGCTACATCTCGGAGCTG 169
      474 CCGGAGCGGCCAAGCTGGCCCTCTGTGCTACCGGCTGCTCATCTTCGCCCTGGGCTC 533
      170 CCGCGCGGGCTAAACTGGCCCTTGGCGCTGGCTGAGGACTCATTTTGGCTGGGCTC 229
      534 TTGGCAATGCTCTGGTGTCTACGTTGTTACCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 593
      230 TTGGCAATGCTCTGGTGTCTATGTTGGTGGACCGGAGGAGGAGGAGGAGGAGGAGGAG 289
      594 AACATCTTTATCTGCTCTTGGCGCTCAGTGACTGCTCATACCTTCTTCTGATTCCTCC 653
      290 AACATCTTATCTGCTCTCTGGCAGCTCAGTGATCTGCTCATCTGCTTCTTCTGATTCCT 349
      654 GTCCATGCTCCAGAACATTTCCGACAACTGGCTGGGGGTGCTTTCATTTGCAAGATG 713
      350 GTCACGATGCTCCAGAACATCTCCGACAACTGGCTGGGTGGTGGTGGTGGTGGTGGTGG 409
      714 GTGCCATTTGTCAGTCTACCGCTGTTGTGACAGAAATCTCTCATATGAGCTGCTGCTGCT 773
      410 GTGCCCTTGTCCAGTCCACTGCTGTGTGTGACGAAATCTTCCATGATCTTGTGATTCGC 468
      774 GTGGAAGGACCCAGGAGCTTGTGATCTCTTTTAAATGAAGTGGCAATACACCAACCGA 833
      469 -TTGAGAGGACCAAGAGACTCATCCATCTTTTAAAGAGAGTGGCAGTACATACCGGG 527
      834 AGGGCTTTCAATGCTAGGTGTGCTGCT -GGTGGCAGTCAATCGTAGGATCACCCA 891
      528 GGGGCTTACAAAATCTTGGGTGTGCTGGTGTGGCAACCATCATCGAGGGGATCACCCC 587
      892 TGTGGCAGCTGC -AACAACTTGAGTCAAAATGATCTTCTTATGA 937
      588 TGGGGACGTACAAACGCTTGAAAATAAGTTTGAATCTTCTCTATGA 634

```

```

RESULT 14
LOCUS   BQ735771                               866 bp mRNA linear EST 16-JUL-2002
DEFINITION
AGENCOURT_8096527 NICHD_XGC_Emb4 Xenopus laevis cDNA clone

```

IMAGE:5571086 5', mRNA sequence.
 B0735771
 VERSION B0735771.1 GI:21874668
 EST.
 KEYWORDS
 SOURCE Xenopus laevis (African clawed frog)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 REFERENCE
 1 (bases 1 to 866)
 NCI-CCAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 TITLE Tumor Gene Index
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Igor Dawid
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: NCI-CCAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
 Plate: L1AM12314 row: f column: 15
 High quality sequence stop: 553.
 FEATURES
 Location/Qualifiers
 1..866
 /organism="Xenopus laevis"
 /mol_type="mRNA"
 /db_xref="taxon:8355"
 /clone="IMAGE:5571086"
 /dev_stage="embryo, stage 31-32"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NICHD XGC_Emb4"
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo
 dt. Average insert size 2.1 kb. Constructed by Life
 Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."
 BASE COUNT 218 a 215 c 195 g 237 t
 ORIGIN
 Query Match 15.4%; Score 370.8; DB 13; Length 866;
 Best Local Similarity 74.2%; Pred. No. 0.52;
 Matches 482; Conservative 0; Mismatches 167; Indels 1; Gaps 1;
 QY 451 CGCTGCTACACCCAGAGTCGGGAGCGGCGGACGCGCTGCTGCTCAGCGGG 510
 DB 18 GTCCGCCACCGGTCCGCCACGCGGTCGCGGCGGCGGCTTTGTGGCAATCTGTG 77
 QY 511 TGCTCATCTTCGCGCTCTTTGGCAATGCTCTGTTGTTTACGTGGTACCGGCA 570
 DB 78 TGCTGATATTGTGTGGCGCTGTTGGCAATCTCTGTGACTTTATGTGTGACCA 137
 QY 571 CGAGGGCCATGGCAGCGCTACCAACATCTTATCTGCTGCTGCGCTCAGTGACCTGC 630
 DB 138 GCAAGCAATGAGGACTGTACCAACATCTTATCTGCTCTCTGGCGCTAGTATCTAC 197
 QY 631 TCATCACCTCTTCTGCAATCCGCTCACCATGCTCCAGAACATTCGCAACCTGGCTGG 690
 DB 198 TCATTGGCTTCTCTGCAATCCCTTCACCATGCTGCGAACATCTCTCCAACTGGCTGG 257
 QY 691 GGGGTGCTTTCATTTGCAAGATGTTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCA 750
 DB 258 GAGGAGGGTTCGCTTGCAAAATGTTTCCCTTTGTTGCTGCTGCTGCTGCTGCTGCTG 317
 QY 751 TCTCCTATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 810
 DB 318 TCTTGCAATGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 377
 QY 811 TGAATGGCAATACACCAACCGAAGGCTTTTCAATGCTAGGTGCTGCTGCTGCTGCTGCTG 870
 DB 378 TGAATGGCAATACACCAACCGAAGGCTTTTCAATGCTAGGTGCTGCTGCTGCTGCTGCTG 437

QY 871 CAGTCATCGTAGGATCACCCATGTGGCAGCTGCAACAACTTGAGATCAATATGACTTCC 930
 DB 438 CAGCAGTGGTGGTATACCAATGTGGCATGCAAGCGCTTAGAGGTAAATATGACTTTT 497
 QY 931 TATATGAAGGAGACACATCTGCTGTAGAGAGTGGACCGCTGTGCACCAAGAA 990
 DB 498 TATATGAAGGAGACACATCTGCTGTAGAGAGTGGACCGCTGTGCACCAAGAA 557
 QY 991 TCTACACCACTTCATCTTGTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1050
 DB 558 TCTATACCGCTTCATCTTGTGTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 617
 QY 1051 TGTACAGT-AAATGCTTATGAACCTTGGATTAAGAAAGAGCTTGGGA 1099
 DB 618 TCTACAGTAAATTTGGCTATGAACCTTGGATTAAGAAAGAGCTTGGGA 667
 RESULT 15
 BU458649
 LOCUS 603367593F1 CSEORBN19 Gallus gallus cdna clone CHEST269123 5', mRNA
 DEFINITION
 ACCESSION BU458649
 VERSION BU458649.1 GI:25947960
 KEYWORDS
 SOURCE Gallus gallus (chicken)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
 Phasianinae; Gallus.
 1 (bases 1 to 888)
 Boardman, P.E., Santa-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
 Fong, W.T., Fickie, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
 A Comprehensive Collection of Chicken CDNAs
 Curr. Biol. 12 (22), 1965-1969 (2002)
 22335534
 12445392
 CONTACT: Simon Hubbard
 Department of Biomolecular Sciences
 University of Manchester Institute of Science and Technology (UMIST)
 PO Box 88, Manchester, M60 1QD, UK
 Tel: 01612008930
 Fax: 01612360409
 Email: Simon.Hubbard@umist.ac.uk.
 Location/Qualifiers
 1..888
 /organism="Gallus gallus"
 /mol_type="mRNA"
 /strain="Layer"
 /db_xref="taxon:9031"
 /clone="CHEST269123"
 /sex="Female"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="CSEORBN19"
 /note="Organ: ovary; Vector: pBluescript II KS(+); Site 1:
 EcoRI; Site 2: NotI; This normalized library was
 constructed from 1 million independent clones. CDNA
 synthesis was initiated using an oligo(dT) primer, using
 methylated C in the first strand synthesis reaction.
 Following this first strand reaction, double-stranded cDNA
 was blunted, ligated to NotI adapters, digested with EcoRI
 , size-selected, and cloned into the NotI and EcoRI
 compatible sites of a custom modified MCS of the
 pBluescript (KS+) vector. The library was normalized in 2
 rounds using conditions adapted from Soares et al., PNAS
 (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
 (1996): 791, except that a significantly longer
 reannealing hybridization was used."
 BASE COUNT 280 a 154 c 187 g 267 t
 ORIGIN

```
Query Match      13.4%; Score 324; DB 13; Length 888;
Best Local Similarity 71.5%; Pred. No. 5.4;
Matches 509; Conservative 0; Mismatches 195; Indels 8; Gaps 6;

QY 922 ATGACTTCCATATATGAAAGGACACATCTGCTGTAGAGAGGACGACCCCTGTGC 981
Db 2 ATGACTTCTGTATGAAAAGTGACATCTG-TGCTTGGAGAAATGGGCCGTCACAACTT 60

QY 982 ACCAGAAGATCTACACCCTTCATCCTTGTCTATCCTCTCTCTCTCTCTCTCTCTATGGTGA 1041
Db 61 ATCAGAAGATATATACCCTTT-ATACCTTGTATATCTCTTTCTTTTCCACTGATATGA 119

QY 1042 TGCCTTATCTGTACAGTAAATTTGGTTATGAACTTTGGATAAGAAAGAGTTGGGGATG 1101
Db 120 TGCCT--TTCTTATACACTAAATGGTTATGAACTCTGGAATTAAGAAACGAGTGGGAGATG 177

QY 1102 GTTCAG-TGCTTCGAACCTATTCATGAAAGAAATGTCACAAATAGCCAGGAGAGAGAA 1160
Db 178 CTTTCAGATCTTCAAAACCATTCATGGGAGTGAATGTCTAAATATCAAGGAAGAGAG 237

QY 1161 CGAGCTGTCATTTATGATGGTGACAGTGTGGCTCTCTTTGTGTGTGTGTGGGCACCATTC 1220
Db 238 CGAGCAAT-GTATGATGGTGACAGTGTGTCTCTTTGTGAGTCTGTGGGCCCTTTC 296

QY 1221 CATGTTGTCATATGATGATGAATACAGTAATTTTGAAGGAATATGATGATCTCA 1280
Db 297 CACATTAATTCATGATGGAATACAGTAATTTTGAAGGAATGATGATGATGATGATG 356

QY 1281 ATCAAGATGATTTTGTCTATCGTCAAAATTTATGGATTTTCCAACTCCATCTGTAATCCC 1340
Db 357 ATCAAAATGATCTTTGCAATTTGTTCAATCATAGGATCTTCAATCTATTTGCAACCTT 416

QY 1341 ATTGTCTATGCAATTTATGAATGAAACTTCAAAAAAATTTTGTCTGCAATTTGTTAT 1400
Db 417 ATTGTGTATGCTTTTCATGAATGAGAACTTCAAGAAAAATTTTCTGTGCGCATCTGCTTC 476

QY 1401 TGCAATGTAATAAATCAACCTTCTCTCCAGCACAAAGGCATGGAATTCAGGAATTCATG 1460
Db 477 TGGCTTGTCAAGAAAAATGCAATCACCAACCGCAACTTGGGAACCTCGGGATTACTATG 536

QY 1461 ATGGGGAAGAAAGCAAAAGTTTTCCTCAGAGAGAAATCCAGTGGAGGA--AACGAAGGAG 1518
Db 537 AGCGGCAAAAGGCGAGGTGATTTCTCAGAGGGCTCCCACTGACTCGGATGAGGCCAGGAGA 596

QY 1519 AAGCAATTCAGTATGGCAACATTTGAAGTCAATTTGTGTAACAGACAGAGGAGAGAAAA 1578
Db 597 GAGGCGTCAGTATGGCAACATTTGAAGTGAAGTTCTGTGACCAACCGTCTTCAAAAAAGGA 656

QY 1579 AGCTCAACGACATCTTCTCTCTTTAGTCTGAACTGGCTGAGAAATCTCC 1630
Db 657 ATTTGAAAGGCACTCTCACTTTATTCAGCTCTGAGCTTCTCGCGCATCTCTGC 708
```

Search completed: October 29, 2003, 11:53:53
Job time : 5007 secs